UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISAS EM SENSORIAMENTO REMOTO E METEOROLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

VALIDAÇÃO DE DADOS GEOGRÁFICOS ESPACIAIS EM AMBIENTE ORIENTADO A OBJETO

MARCUS FABIANO SILVA SALDANHA
Engenheiro Militar – Cartógrafo
Orientador: Prof. Dr. Francisco Humberto Simões Magro

Porto Alegre (RS), outubro de 2005.

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Sensoriamento Remoto, área de concentração Geoprocessamento.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISAS EM SENSORIAMENTO REMOTO E METEOROLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

VALIDAÇÃO DE DADOS GEOGRÁFICOS ESPACIAIS EM AMBIENTE ORIENTADO A OBJETO

MARCUS FABIANO SILVA SALDANHA
Engenheiro Militar – Cartógrafo

Dissertação apresentada ao Curso de Pós-Graduação, em Sensoriamento Remoto e Meteorologia da Universidade Federal do Rio Grande do Sul, como requisito parcial à obtenção do título de mestre.
Orientador: Prof. Dr. Francisco Humberto Simões Magro

Porto Alegre
Outubro de 2005

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISAS EM SENSORIAMENTO REMOTO E METEOROLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

VALIDAÇÃO DE DADOS GEOGRÁFICOS ESPACIAIS EM AMBIENTE ORIENTADOA A OBJETO

MARCUS FABIANO SILVA SALDANHA
Engenheiro Cartógrafo – UFPR
Engenheiro Militar – IME

Esta dissertação foi julgada para obtenção do título de MESTRE EM SENSORIAMENTO REMOTO e aprovada em sua forma final pelo Programa de Pós-Graduação em Sensoriamento Remoto.

Prof. Dr. Franscisco Humberto Simões Magro
CEPSRM - Orientador

Prof. Dr. Sérgio Florêncio de Souza
Coordenador do CPGSR

Prof. Dr. Jorge Ricardo Ducati
Diretor do CEPSRM

COMISSÃO EXAMINADORA

Prof. Dr. Norberto Dani
UFRGS

Prof. Dr. Ronaldo dos Santos da Rocha
UFRGS

Prof. Dr. Sérgio Florêncio de Souza
Coordenador do CPGSR

OFERECIMENTO

Aos meus pais, meus primeiros educadores, que me encaminharam no aprendizado da vida, a minha eterna companheira Paula Akemi e a nossa filha Larissa Yumi, com as quais divido a satisfação pela realização deste trabalho.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Magro, pela confiança depositada na idéia que originou este trabalho.

A Diretoria de Serviço Geográfico do Exército Brasileiro, personificada pelo Excelentíssimo General de Brigada Paiva de Sá, por ter me concedido a oportunidade de realizar o curso de mestrado.

Ao Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia, personificado pelo seu corpo docente, por ter me repassado os conhecimentos necessários para fundamentar o desenvolvimento do trabalho de dissertação.

A 1ª Divisão de Levantamento do Exército Brasileiro, personificado pelo seu corpo técnico por ter compartilhado idéias, informações e camaradagem, durante a elaboração das soluções e experimentos apresentados neste trabalho.

A Deus... por ter permitido nossa existência.

LISTA DE FIGURAS, QUADROS E TABELAS

FIGURAS

Figura 2.1 - Modelo de transferência de informação cartográfica, com exemplos de mapas permanente, virtual, temporal e mental 14

Figura 2.2 - As funções dos gráficos numa seqüência de pesquisa 16

Figura 2.3 - O “cubo cartográfico” 20

Figura 2.4 - A natureza dos dados espaciais: da realidade (a), via construção de modelo e seleção para um modelo de paisagem digital (b), seguido pela seleção e construção para uma representação cartográfica que dá origem a um modelo cartográfico digital (c), apresentado como um mapa (d), o qual resulta no mapa mental do usuário (e) 22

Figura 2.5 - Atributos armazenados num SIG 24

Figura 2.6 - Poder de cruzamento de informação num SIG: sobreposição de dados vetor e raster 25

Figura 2.7 - A visão do SIG como metodologia integradora 26

Figura 2.8 - Os principais componentes de um SIG 28

Figura 2.9 - Componentes de um sistema de informações geográficas 28

Figura 2.10 - Importância da cartografia automatizada para alguns tipos de sistemas de informação espacial 31

Figura 2.11 - Relações interdisciplinares entre SIG e outras áreas 32

Figura 2.12 - Definição técnica de geoprocessamento 33

Figura 2.13- O conceito de camadas ou *overlay* do mundo real 39

Figura 2.14 - Representação de feições geográficas no ambiente digital: ponto (a), linha (b) e polígono (c). Representação de feições de área e situações mais comuns: áreas disjuntas (d), áreas sobrepostas (e) e as duas situações ao mesmo tempo (f) 40

Figura 2.15 - Representação de uma linha (a) em estrutura vetorial (b) e em estrutura raster (c) 40

Figura 2.16 - Modelo "espaguete" de dados 42

Figura 2.17 - Armazenamento de dados de pontos: exemplo para feição calssificada como escola 44

Figura 2.18 - Estrutura de dados para rede 44

Figura 2.19 - Representação de grafos não-direcionados 45

Figura 2.20 -Relacionamento de polígonos adjacentes 46

Figura 2.21 Extrato da TBCD indicando os atributos usados para codificar a feição cartográfica para o modelo digital 57

Figura 2.22 - Detalhe da estrutura matricial com resolução de 1 bit, princípio da formação do espaço "booleano" 72

Figura 2.23 - Erro em área para a representação de um quadrado de 20 x 20 mm numa escala 1:250.000 89

Figura 2.24 - Digitalização vetorial de uma linha curva 90

Figura 2.25 - “Onde o ponto xy se encontra em relação ao polígono A?”: (1) definitivamente contido em A; (2) definitivamente não contido em A; (3) possivelmente contido em A; (4) possivelmente não contido em A; (5) sobre a linha limite de A digitalizada em forma de vetor 92

Figura 2.26 - Erros potenciais do processo de vetorização 94

Figura 2.27 - Lógica do algoritmo Douglas Peucker, para eliminação de vértices em excesso 98

Figura 2.28 - Características de duas classes representadas graficamente por retângulo dividido em três partes: a superior contém o nome da classe; a do meio contém os atributos e a inferior contém os métodos definidos pelo usuário 104

Figura 2.29 - Hierarquia e herança. Criação da classe Veículo para conter as classes Caminhão e Ônibus. Somente as características próprias de cada subclasse são mantidas na mesma 104

Figura 2.30 - Exemplo de herança múltipla para ao caso da classe Ônibus e herança simples para a classe Caminhão 105

Figura 2.31 - Herança múltipla 106

Figura 2.32 - Temas da modelagem do Espaço Geográfico Brasileiro 109

Figura 2.33 - Estrutura básica dos Temas da modelagem do espaço Geográfico Brasileiro 110

Figura 2.34 - Arquitetura do sistema *Gothic* 111

Figura 2.35 - Conectividade entre elementos lineares 121

Figura 2.36 - Captura de curso d´água visando manter a integridade do sistema de rede formado 121

Figura 2.37 - Formação de feição de área definida a partir de feições lineares 122

Figura 2.38 - Visão do usuário e dados armazenados no banco de dados 124

Figura 2.39 - Conexão entre link e nó para definir a conectividade entre os objetos 124

Figura 2.40 - Árvore de versões simples. Fonte: Laser-Scan, 2000 127

Figura 2.41 - Criação de *dataset* filho para aplicação militar 129

Figura 3.1 - Fluxo de geral produção de dados geográficos no âmbito da DSG, situando a validação e estruturação de dados vetoriais 143

Figura 3.2 - Fluxo de validação e estruturação de arquivos vetoriais 144

Figura 3.3 - Exemplo de feição que deve ser estendida na categoria limites 153

Figura 3.4 - Exemplo de feição que deve ser estendida na categoria transportes 153

Figura 3.5 - Exemplo de falsa ponta livre na categoria hidrografia 154

Figura 3.6 - Exemplo de linha pequena que pode acarretar erro em sistemas de rede.....154

Figura 3.7 - Exemplo de pequeno segmento 155

Figura 3.8 - Exemplo de ângulo incorreto 155

Figura 3.9 - Exemplo de elementos próximos e incorretos 156

Figura 3.10 - Exemplo de interseção incorreta 156

Figura 3.11 - Polígono de Vegetação limitando pela estrada 157

Figura 3.12 - Área de vegetação delimitada pela hidrografia 157

Figura 3.13 - Janela da rotina FLINHAC 162

Figura 3.14 - Janela da rotina FSHAPED 162

Figura 4.1 - Regra para eliminação de elementos adotada na rotina EliminarDup 174

Figura 4.2 - Linha desenhada com *loops* antes e após a aplicação da rotina AbrirLoops.....174

Figura 4.3 - Aplicação da rotina ExtLinhasFlut em *dataset* na escala 1:25.000 175

Figura 4.4 - Aplicação da rotina ExtLinhasFlut sobre elementos menores que 30 m 175

Figura 4.5 - Situação possível de alguns elementos lineares antes (a) e após a aplicação da rotina ExtLinFlut (b) e o *loop* gerado 176

Figura 4.6 - Aplicação da rotina GerarIntersecoes em cruzamentos de elementos lineares 176

Figura 4.7 - Situação nas quais o segmento de linha curto será eliminado.....177

Figura 4.8 - Possível elemento duplicado: (a) gerado após aplicação das rotinas ExtLinhasFlut e GerarIntersecoes e (b) detectado para eliminação pela rotina EliminaDup.....177

Figura 4.9 - Elementos sobrepostos após a execução de EliminaDup podem configurar segmentos curtos 178

Figura 4.10 – Marcação dos vértices livres para serem checados pelo operador.....179

Figura 4.11 - Marcação de vértices que originam ângulos pequenos 180

Figura 4.12 - Situação marcadas pela rotina FlagInt/Loop 180

Figura 4.13 - Situação na qual o elemento de área está aberto e apresenta um vértice livre.....181

Figura 4.14 - Buracos existentes num objeto de área tipo rio de margem dupla 184

Figura 4.15 - Inserção de centróide de vegetação de forma generalizada sobre curso d´água.....186

Figura 4.16 - Classificação de área não classificada por consulta de adjacência 186

Figura 4.17 - Fluxo de trabalho atual na 1ª DL 192

Figura 4.18 - Fluxo de trabalho proposto 192

QUADROS

Quadro 2.1 - Fontes de possíveis erros para os SIG 87

Quadro 2.2 - Código convencional para a execução de operação display que recebe um objeto como entrada e o apresenta na tela 107

Quadro 3.1 - Níveis de validação primária e secundária de um banco de dados 135

TABELAS

Tabela 2.1 Relação entre feições geográficas, entidades do AutoCAD, conteúdos do Banco de Dados e feições de SIG 51

Tabela 2.2 - Atributos existentes no *software* MicroStation 55

Tabela 2.3 - Fatores de conversão adotados para dimensionamento dos simbolos cartográficos contidos na bibilioteca de células 56

Tabe 2.4 Especificaçoes técnicas para a digitalização matricial de originais cartograficos na DSG 63

Tabela 2.5 - Precisões estabelecidas para o georreferenciamento de cartas nas escalas do mapeamento sistemático 65

Tabela 2.6 - Valores de *gap* para vetorização de linhas matriciais descontínuas 78

Tabela 2.7 - Valores de ampliaçao de escala para as células usadas pela DSG 80

Tabela 2.8 - Regras de topologia existentes no *Gothic* (iterações possíveis entre primitivas).....126

Tabela 3.1 - Datuns e arquivos semente criados pela DSG para a plataforma de trabalho MGE 150

Tabela 3.3 - Arquivos de categorias e de textos associados 152

Tabela 4.1 - Seqüência de temas da TBCD a serem percorridos na Fase I 172

Tabela 4.2 - Tolerâncias da rotina FlagPeqSeg 179

Tabela 4.3 - Objetos delimitadores de área 183

SUMÁRIO

Capítulo 1 – INTRODUÇÃO 1

1.1 Identificação do problema 1

1.2 Definição do objetivo geral e específico do trabalho 5

1.3 Estrutura do trabalho 7

Capítulo 2 – CONCEITUAÇÃO TEÓRICA 10

Parte I – CARTOGRAFIA AUTOMATIZADA E SISTEMAS DE INFORMAÇÃO GEOGRÁFICA 10

2.1 Sistemas de informação geográfica e cartografia 10

2.1.1 Mapas para visualizar o mundo 11

2.1.2 Cartografia: Definição, Objeto e Objetivo 12

2.1.3 Processo de comunicação cartográfica 15

2.1.4 Dados espaciais 20

2.1.5 Sistemas de Informações Geográficas 23

2.1.5.1 Definição 23

2.1.5.2 Componentes de um SIG 26

2.1.6 Geotecnologias 29

2.1.7 Relação entre SIG e cartografia 34

Parte II - AQUISIÇÃO DE DADOS GEOGRÁFICOS ARMAZENADOS EM ESTRUTURA VETORIAL 38

2.2 Aquisição de dados 38

2.2.1 Entrada de dados no SIG 38

2.2.2 Estrutura de dado usada na cartografia digital 39

2.2.3 Modelo de estrutura de dados vetorial 41
Estrutura de dados para armazenar Pontos 43
Estrutura de dados para armazenar Linhas 43
Estrutura de Dados para Armazenar Polígonos 45
2.2.4 A confecção de mapas com CAD 47
2.2.4.1 Diferenças entre CAD e SIG 47
Criação de topologia 47
Sistemas de coordenadas 48
Ligação a atributos 48
Outras 48
2.2.4.2 O Registro de dados no CAD 49
2.2.4.3 Relacionamento das entidades do CAD com as feições do SIG 50
Entidades CAD x tipos de feições SIG 50
Organização do dado CAD 52
2.2.4.4 Conversão de CAD para SIG 52
Leitura direta de dados 52
Tradução a partir do formato DXF 53
Programas híbridos CAD/SIG 53
2.2.5 Estruturação de dados vetoriais em CAD 54
2.2.6 Aquisição de dados vetoriais para SIG via CAD 58
2.2.6.1 Processo de digitalização vetorial ou vetorização 58
Digitalização matricial 58
Tipo de originais 60
Requisitos de Qualidade 60
Preparo 64

Confecção de moldura 64

Registro 64

Originais Cartográficos 65

Carta Impressa 70

Vetorização 71

Vetorização em tela 71

Vetorização de originais cartográficos 71

Vetorização de cartas impressas 73

Vetorização manual (com mesa digitalizadora) 75

Aspectos importantes 76

Procedimentos adotados na vetorização 79

Parte III – QUALIDADE DE DADOS EM ESTRUTURA VETORIAL PARA USO EM SIG 83

2.3 Qualidade de dados para SIG 83

2.3.1 Entraves ao uso do SIG 85

2.3.2 Fontes de erros dos dados 86

2.3.3 Erros associados à digitalização vetorial de uma carta 88

Erros associados à origem dos dados 88

Erros associados à representação digital dos dados 89

Erros associados à sobreposição de camadas de elementos de área 92

Erros comuns ocasionados durante a aquisição de dados 93

Vértice após o ponto (*overshot*) 94

Vértice formando pico 94

Área sem centróide 94

Ausência de elemento 95

Vértice antes do ponto (*undershot*) 95

Área aberta 95

Elemento duplicado 95

2.3.4 Algoritmo Douglas-Peucker 96

Parte IV – O AMBIENTE DE BANCO DE DADOS ORIENTADO A OBJETO 99

2.4 Banco de Dados Orientado a Objeto 99

2.4.1. Conceitos Básicos 100

Objetos e Identidade 100

Valores 101

Estrutura do objeto 101

OIDs x Chaves Primárias 102

Objetos Complexos 102

Encapsulamento 102

Métodos 103

Tipos e Classes 103

Herança 103

Polimorfismo 106

2.4.2 Modelagem de dados orientada a objetos para fins cartográficos 107

Modelagem de dados da DSG 108

2.4.3 Sistema *Gothic* 109

2.4.3.1 Arquitetura do sistema 111

Interface do Sistema Operacional 111

X-Server 112

Data Access Manager 112

Interface de Programação de Aplicação 112
Aplicações 113
Banco de Dados 113
Personalização 113
2.4.3.2 Conceitos fundamentais do sistema 114
DataBase e *Datasets* 114
Conjunto de Dados (*Datasets*) 114
2.4.3.2.1 Tipos de Conjuntos de Dados (*dataset*) 115
Dataset do Mundo Real 115
Dataset de Apoio 116
2.4.3.3. Banco de Dados Orientado a Objetos *Gothic* 116
Modelo de Dados Objeto 116
Esquema 117
Classes 117
Objetos 118
Valores 118
Referências 118
Métodos e Comportamentos 118
Composição de Objetos 118
Herança 118
Temas 119
2.4.3.4 Topologia no BDOO do *Gothic* 120
Necessidade do uso da topologia 120
Conectividade 120
Captura 121

Formação de polígonos 121
Tipo de Estruturas de conjuntos de dados 122
Tipo *Spaghetti* 122
Tipo *Link-Node* 122
Classes topológicas básicas 122
Graphic 123
Simple 123
Links e Nós 123
Armazenamento *link-nó* 124
Conexão entre link e nó 124
Topologia *on the fly* 125
Regras de topologia 125
2.4.3.5 Conceitos Básicos de Versionamento do BD *Gothic* 125
Pontos de Checagem e de Retorno (*Checkpoints* e *Backtracking*) 127
Compressing 128
Conjunto de dados Filho (*Childs*) 129
Acesso a múltiplos usuários 129
Edição por múltiplos usuários 130
Transações Longas 130
Merging 132
Clones 133

Capítulo 3 – METODOLOGIA 134
3.1 Validação de banco de dados geográficos 134
3.1.1 Propósito 134

3.1.2 Responsáveis pelo processo 136
3.1.3 Necessidade de validar os dados 136
a. Dado digital 136
b. Aumento de compartilhamento e produção descentralizada de dados 137
c. Maior custo de produção 137
3.2 Estruturação e validação de dados vetoriais para SIG 138
3.2.1 Metodologia e Objetivo 139
3.2.2 Critérios para criação de classes de elementos espaciais 139
3.2.3 Definição das categorias 140
3.2.4 Fluxo geral de dados 143
3.2.5 Fluxo de validação e estruturação 143
3.2.6 Rotinas utilizadas 144
MDL (*MicroStation Development Language*) 144
MACRO 147
3.2.7 Descrição das fases da validação 148
3.2.7.1 Preparo para validação 148
3.2.7.2 Conexão ODBC 149
3.2.7.3 Configuração da plataforma de trabalho 149
3.2.7.4 Preparo do arquivo geral 150
3.2.7.5 Validação das categorias 151
Seqüências de operações para validação das categorias 152
a. Todas as categorias 152
b. Categoria vegetação 157
3.2.7.6 Revisão e correção 160
Impressão 160

Revisão 160

Correção 161

3.2.7.7 Vínculo com o banco de dados 161

Vinculo de elementos lineares e pontuais 161

Vinculo de elementos de área 162

União de arquivos de linhas, pontos e de áreas 163

Verificação final dos arquivos 163

Capítulo 4 – RESULTADOS 164

4.1 Validação em Ambiente Orientado a Objetos 164

4.2 Validação de Dados no ambiente Orientado a Objetos 165

4.2.1 Implementação das rotinas de validação 166

4.2.2 Linguagem LULL 166

4.2.2.1 Funções e Bibliotecas em LULL 167

4.2.2.2 Programação em LULL 167

Comentário*s* 167

Variáveis 168

Declaração de variáveis 168

If Statements 168

Loops 168

4.2.3 Elaboração das rotinas de validação 169

Estudo das rotinas usadas no CAD 169

Estudo da linguagem de programação LULL 169

Analise das funções existentes na biblioteca do *Gothic* 170

Estudo da Modelagem de dados do Espaço Geográfico Brasileiro 170

Escrita dos códigos na linguagem LULL 171
Testes com o aplicativo 171
4.3 Metodologia de Validação no ambiente OO 171
4.3.1 Fase I 172
Rotinas da Fase I 173
Etapa Automática 173
Etapa Iterativa 179
4.3.2 Fase II 181
Rotina da Fase II 182
4.3.3 Fase III 183
Rotinas da Fase III 184
Criação de objeto de área a partir do atributo cor 184
Criação de objeto de área a partir do centróide 185
4.3.4 Fase VI 187
4.4 Avaliação da metodologia 187
4.4.1 Resultados Obtidos 188
4.4.2 Considerações Finais 190

Capítulo 5 – CONCLUSÕES E RECOMENDAÇÕES 194
5.1 Conclusões 194
5.2 Recomendações 196

REFERÊNCIAS BIBLIOGRÁFICAS 197

VALIDAÇÃO DE DADOS GEOGRÁFICOS ESPACIAIS EM AMBIENTE ORIENTADO A OBJETO

Autor: Marcus Fabiano Silva Saldanha
Orientador: Prof. Dr. Francisco Humberto Simões Magro

RESUMO

A produção de base cartográfica para os SIG apresenta peculiaridades distintas daquelas produzidas somente para atender os usuários da cartografia. Portanto cartas digitais em formato vetor produzidos em ambiente de CAD devem passar por rigoroso processo de validação geométrica antes de serem usadas por tais sistemas. Este trabalho tem por finalidade apresentar uma metodologia de validação geométrica de dados geográficos para uso em SIG baseada em ambiente orientado a objeto. Pra tanto são apresentados no decorrer do trabalho os principais conceitos relacionados ao assunto, bem como, a descrição da aquisição de dados via digitalização vetorial, as principais fontes de erro oriundas deste processo e o processo de validação em ambiente de CAD. Também é apresentado de forma pormenorizada os conceitos relacionados à tecnologia de banco de dados Orientado a Objeto e o Sistema *Gothic*, dentro do qual foram realizadas as implementações. Por fim são apresentadas a descrição das rotinas elaboradas e a metodologia desenvolvida, além de resultados obtidos em testes preliminares. Os testes realizados mostram que ocorreram ganhos de produção consideráveis sem que houvesse perda de qualidade.

Dissertação de Mestrado em Sensoriamento Remoto, Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia, Curso de Pós Graduação em Sensoriamento Remoto da Universidade Federal do Rio Grande do Sul.
Porto Alegre - RS, 03 de outubro de 2005.

VALIDATION OF ESPACIAL GEOGRAPHIC DATA IN ORIENTADED OBJECT ENVIRONMENT

Author: Marcus Fabiano Silva Saldanha
Orientation: Prof. Dr. Francisco Humberto Simões Magro

ABSTRACT

The production of cartographic base for the SIG presents distinct peculiarities of those only produced to attend the cartographic users. Therefore digital charts produced in vector format in CAD environment must pass for rigorous process of geometric validation before being used for such systems. This work has for purpose to present a methodology of geomettric validation of geographic data for use in SIG based on oriented object environment. For this the main concepts related to the subject are presented during the work, as well as, the description of acquisition of the data by means of vector digitizer, the main deriving sources of error of this process and the process of validation in CAD environment. Also it is presented of detailed form the concepts related to the technology of Oriented Object data base and the Gothic system, which was used for the implementations. Finally the description of the elaborated routines and the developed methodology are presented, beyond some results gotten in preliminary tests. The tests show that considerable profits of prodution had occurred without that had loss of quality.

Dissertation of Master in Remote Sensing, State Center of Researches in Remote Sensing and Meteorology, Course of Degree in Remote Sensisng of the Federal Univetrsity of Rio Grande do Sul, Brazil.
Porto Alegre - RS, October 03, 2005.

Capítulo 1 – INTRODUÇÃO

1.1 Identificação do problema

A tendência de gerenciamento de informação por meio computacional é hoje uma realidade e a sua dinâmica um processo irreversível. A evolução da micro-informática nos últimos anos contribuiu significativamente para este panorama, criando uma dependência parcial da humanidade por tal tecnologia. Os microcomputadores fazem parte do dia a dia das pessoas e são considerados ferramentas fundamentais no ambiente de trabalho das empresas. A integração de dados por meio de sistemas de redes a nível local e global, por sua vez, também é uma realidade e a cada ano, assim como a micro informática, apresenta inovações que são prontamente absorvidas pela sociedade.

Mapas, por sua vez, têm sido usados por séculos para visualizar dados espaciais. Além disso, a coleta de dados distribuídos espacialmente na superfície terrestre tem sido uma atividade constante das sociedades organizadas. Das mais antigas civilizações aos tempos modernos, a coleta e tratamento para transformar esse tipo de dado em informação visual útil vem acompanhando o próprio progresso da humanidade. Originalmente os mapas eram usados para descrever lugares distantes, como um aliado para navegação e para estratégias militares [Hodgkiss 1981]. Hoje desempenham um papel que vai muito além das propostas originais auxiliando desde o cadastro de imóveis urbanos nas cidades até o estudo para implantação de redes logísticas de distribuição e atendimento de grandes empresas a níveis internacionais.

O grande estímulo para atingir esse nível de aplicabilidade, porém ocorreu somente na década de 80, com o desenvolvimento da informática em termos de hardware e de software. Durante esse período o desenvolvimento tecnológico voltado para o manuseio de dados espaciais em ambiente digital foi relativamente alto, fazendo com que a maioria dos empreendimentos onde era necessário o uso de mapas sofressem alterações significativas. O desenvolvimento também possibilitou o aprimoramento e a popularização de *softwares* de computação gráfica que impulsionados pelas necessidades de obtenção de informações a partir de dados espaciais, ampliou consideravelmente o universo de usuários

e produtores de informações geográficas. A união dos mapas digitais com sistemas de Banco de Dados, por sua vez permitiu realizar pesquisas consistentes sobre dados geográficos fornecendo um terreno fértil para o florescimento de novas ferramentas tecnológicas.

Os pacotes de softwares que possibilitavam consultas e análises de dados espaciais deste tipo ficaram conhecidos como Sistemas de Informação Geográfica (SIG). A disseminação desta ferramenta, que até pouco tempo era usada somente por instituições públicas em países mais desenvolvidos, e mesmo assim sem explorar sua total capacidade, fez com que sua funcionalidade amadurece-se proporcionando sua aplicabilidade para todas as disciplinas que trabalhavam com dados geográficos, possibilitando pela primeira vez a integração dos dados de diferentes tipos de fonte.

Hoje sua funcionalidade possibilita manipular, analisar e visualizar dados combinados de várias naturezas, tornando-o uma ferramenta imprescindível para qualquer usuário que lance mão de dados espaciais. O SIG permite aos usuários, através da ligação de seus modelos baseados em aplicativos com dados contidos no sistema, tentar encontrar respostas para questões tais como “qual o local mais adequado para implantar uma nova sede para uma rede de supermercados?”, ou “qual será o efeito deste empreendimento na área ao redor do local implantado?”.

A ampliação do número de usuários de SIG tem desempenhado um papel importante e positivo para o amadurecimento da cultura cartográfica, pois ao gerar a demanda por informações espaciais com qualidade, desencadeiam processos que conseqüentemente determinam a política a ser adotada para a produção das bases cartográficas. Dessa forma o usuário contribui de maneira significativa para o aprimoramento técnico dos núcleos produtores daquelas informações, sejam eles públicos ou privados. Dentro deste contexto, a otimização do processo de seleção das fontes e dos métodos de aquisição de dados, assim como, a conversão desses dados para a forma digital impõe a necessidade de adoção de normas e procedimentos que garantam a qualidade das bases de dados gráficos.

Cabe ressaltar que para um SIG possa fornecer informações confiáveis como apoio à tomada de decisão, a consistência lógica dos dados espaciais deve estar preservada e isenta de erros. As fontes de erros nesse tipo de sistema podem ter duas origens: uma nos

dados alfanuméricos inseridos no banco de dados utilizado, estando os erros relacionados à coleta ou cadastro desse tipo de informação no sistema, a outra diz respeito aos dados espaciais, ou seja, aos elementos gráficos que representam espacialmente a distribuição das informações. Estes erros estão relacionados ao tratamento inadequado após aquisição ou ainda a fontes de dados não confiáveis.

Para eliminar, ou em última análise minimizar, os erros citados é necessária a adoção de ações preventivas na fase responsável pela fonte de erro. Estas ações devem ser baseadas em procedimentos técnicos que devem ser rigidamente executados, visando como meta final o controle de qualidade das informações.

Para o caso de inconsistências oriundas do cadastro incorreto de dado alfanumérico nos bancos de dados os procedimentos de controle e correção dependem diretamente de ações humanas, pois a natureza das informações em questão não permite o emprego de outros recursos. Este tipo de informação, porém geralmente é cadastrado e gerenciado por órgãos e instituições públicas, o que de certa forma é uma garantia para a qualidade e validade das informações.

Para o outro caso, ou seja, a informação espacial cujas inconsistências têm origem no tratamento inadequado após aquisição é possível o uso de recursos computacionais no auxílio à correção dos dados. Esta tarefa pode ser realizada utilizando rotinas pré-existentes na maioria dos sistemas disponíveis no mercado ou através da implementação dessas rotinas em linguagem de programação adequada. Além das instituições públicas, algumas empresas privadas do ramo de mapeamento são responsáveis pela confecção e validação desse tipo de dado.

A cartografia automatizada, encarada como a tecnologia encarregada de realizar a aquisição e validação, compreende os processos de produção de cartas e mapas a partir de dados espaciais armazenados no formato numérico digital ao invés de gráfico analógico. Enquanto o conceito de mapeamento digital é simples na prática sua implantação se torna um pouco mais complexa. O processo produtivo de cartografia em ambiente digital possibilita a implantação de diversos métodos de trabalho. Estes, por sua vez precisam ser estruturados de maneira a distribuir as tarefas em fases de produção como numa linha de montagem. Cada fase é responsável pela execução de certos números de procedimentos e verificações. O resultado final das fases de produção é o produto cartográfico impresso ou

em meio digital conforme o objetivo do projeto executado. Assim como qualquer tecnologia inovadora, os avanços obtidos com o advento da computação permitiram aprimorar processos e solucionar diversos problemas encontrados nas técnicas anteriores. Entretanto estes avanços também foram seguidos de novos desafios a serem superados, que se apresentam principalmente em forma de estruturação metodológica de atividades. O problema maior da cadeia produtiva neste ambiente, portanto não é a tecnologia em si ou o ambiente propriamente dito, mas sim a padronização minuciosa de procedimentos que garanta ao final do processo produtivo um material cartográfico confiável e de qualidade. De forma sintética as fases de produção atual usando os meios digitais se resumem em: aquisição, tratamento e simbolização de dados.

A produção de base cartográfica para atender os SIG apresenta algumas características específicas diferentes da cartografia automatizada, pois nessa, além da apresentação cartográfica, passam a estar envolvidos como objetivos prioritários, a manutenção da integridade de atributos lógicos e o perfeito relacionamento das diferentes feições representadas nas bases de dados gráficos, no sentido de prover ao usuário o acesso ágil e seguro à informação georreferenciada. A conseqüência é que os procedimentos técnicos e os recursos humanos envolvidos nesse processo necessitam de aprimoramento qualitativo a fim de manter o elevado grau de qualidade necessário.

A cartografia digital tem na ciência da computação um instrumento imprescindível, sendo esta amplamente empregada nas diversas etapas da construção das bases de dados. A informática se faz presente desde a aquisição do dado primário por meio dos levantamentos até a disponibilidade final da informação em dispositivos de impressão e visualização. A necessidade de padronização de procedimentos e o desenvolvimento de metodologias voltadas à otimização visando controle de qualidade de produtos em meio digital são, portanto condições fundamentais para a produção de produtos cartográficos em ambiente digital.

Dessa forma, o desenvolvimento de rotinas que permitam a correção de dados espaciais e a agilização do processo de produção cartográfica em ambiente digital é um fator extremamente importante. Por meio disso pode-se garantir a integridade lógica das feições cartográficas o que é fundamental para atingir o padrão de qualidade necessário para utilização dos dados em SIG.

Dentro desse panorama, os sistemas baseados em tecnologia de ambiente Orientado a Objeto (OO), tornam-se aliados naturais para a melhoria das atividades de produção cartográfica, pois permitem que as feições sejam representadas por objetos com comportamentos modelados de acordo como são conhecidos na realidade. Esta vantagem é impar entre os sistemas existentes, permitindo maior liberdade no manuseio, tratamento e análise de informações geográficas. A implementação de procedimentos e rotinas dentro desse ambiente, visando o tratamento após aquisição de dados para SIG, pode ser um avanço em direção à conquista da eficácia e qualidade produtiva de dados espaciais.

1.2 Definição do objetivo geral e específico do trabalho

No Brasil a Diretoria de Serviço Geográfico (DSG) do Exército Brasileiro juntamente com o Instituto Brasileiro de Geografia e Estatística (IBGE), divide a responsabilidade pela confecção de cartas do mapeamento sistemático nacional. Estas instituições até o final da década de 80 produziam seu material cartográfico apenas em meio analógico através da impressão em papel de cartas topográficas. Esses documentos cartográficos, ainda hoje são a principal fonte de dados geográficos para os SIG que trabalham com dados compatíveis as escalas preconizadas para o mapeamento sistemático.

A partir de meados da década de 90 a DSG seguindo os avanços tecnológicos da época realizou a modernização de seus equipamentos restituidores analógicos por meio da adaptação de dispositivo tipo *encolder* (conversor analógico digital) ligados a computadores com sistemas de CAD (*Computer Aided Design* – Desenho Assistido por Computador). Isto possibilitou a produção de dados espaciais em meio digital por meio do processo de digitalização vetorial. Num primeiro momento esses dados não recebiam qualquer tipo de tratamento após a aquisição, pois a necessidade de dados espaciais para SIG não era tão significativa. No final da década de 90, porém houve uma crescente procura por dados tratados para SIG. A diferença deste tipo de dados, quando comparados aos oriundos das primeiras cartas digitais, está na ausência de inconsistências lógicas e erros geométricos, os quais prejudicam o bom funcionamento do SIG. A DSG a partir do ano 2000, pressionada pela grande demanda, implementou em sua linha de produção uma fase responsável pelo preparo de dados com a finalidade de atender aos usuários SIG. Esta

fase recebeu a denominação dentro da instituição de Validação e tem por finalidade realizar a edição para SIG dos dados vetor obtido pelas fases de aquisição. A partir daquele ano todos os arquivos digitais de cartas existentes ou em fase de confecção passaram a ser validados. Hoje a validação se encontra bem sedimentada como processo produtivo e faz parte da linha de produção de dados cartográficos da DSG.

Ainda no ano de 2000, a DSG por intermédio da 1ª Divisão de Levantamento (1ªDL), adquiriu da empresa Laser-Scan Ltda ferramentas de "geotecnologias" baseadas em ambiente Orientado a Objeto (OO). Essas consistem de um sistema de tratamento, manipulação e armazenamento de dados geo-espaciais, conhecido como LAMPS2 e um Banco de Dados Orientado a Objeto (BDOO), denominado *Gothic*. Durante o período entre os anos de 2000 e 2002 essas tecnologias foram implementadas para uso na linha de produção cartográfica da 1ªDL, a qual apresentou excelentes resultados em termos de redução de tempo e aumento de qualidade em processos de edição vetorial. O potencial das tecnologias adquiridas, porém não se resume apenas a implementação de ferramentas para uso em apenas uma fase do processo de produção cartográfica. Ainda existem possibilidades de otimizações de processos em fases de produção ligadas a aquisição de dados, via restituição fotogramétrica digital (LAMPS2 / SOCET SET) ou digitalização vetorial de originais cartográficos (VTRAK) e tratamento destes dados para SIG. Essa última atividade, particularmente, amplia as perspectivas de disponibilidade de dados espaciais estruturados e tratados para SIG o que possibilita, devido à redução de tempo de produção e aumento qualidade, atender a demanda instaurada nos últimos anos.

Dentro deste panorama se faz necessário realizar estudos visando à elaboração de procedimentos de validação dentro do sistema *Gothic*. Este trabalho de dissertação trata basicamente deste assunto e tem por objetivo fundamental contribuir de alguma forma para a evolução da automatização da cartografia dentro do âmbito da DSG.

O objetivo é desenvolver rotinas computacionais e procedimentos operacionais confiáveis e adequados para possíveis aplicações nas linhas de produção das Organizações Militares Diretamente Subordinadas (OMDS) a DSG. Para atingir o objetivo foram estabelecidos basicamente os objetivos especificados a seguir.

a. desenvolvimento das rotinas e metodologia visando a otimização da validação dos dados geográficos em formato vetor para uso em Sistemas de Informações Geográficas (SIG) dentro do ambiente Orientado a Objeto (OO);

b. avaliação preliminar do resultado da aplicação das ferramentas desenvolvidas de maneira a verificar as potencialidades reais do ambiente OO nas atividade de tratamento de dados vetor; e,

c. verificar as possíveis alterações causadas pela implantação da metodologia proposta na linha de produção atual da DSG e suas conseqüências.

Para tanto o trabalho de pesquisa relatado nesta dissertação tomou por base a experiência técnica existente na DSG bem como referências bibliográficas existentes atualmente.

1.3 Estrutura do trabalho

Para atingir os objetivos traçados o trabalho foi elaborado em forma de capítulos os quais foram estruturados de maneira a abordar os diversos aspectos relativos ao assunto.

O Capitulo 2 que trata da conceituação teórica é composto por quatro temas. O primeiro aborda basicamente os aspectos relativos a Cartografia e aos SIG. Ele apresenta conceitos importantes para o entendimento da Cartografia como ciência, bem como realiza uma abordagem sobre os dados espaciais de maneira a inseri-los dentro do contexto dos SIG. Também apresenta os principais conceitos relacionados ao SIG, sua natureza e principais componentes. Aborda ainda o conceito das “geotecnologias” e apresenta uma discussão entre a cartografia e SIG.

O segundo tema trata do aspecto da aquisição de dados geográficos e suas peculiaridades. No seu corpo aborda os aspectos de entrada de dados no SIG e expõe de maneira mais abrangente a estrutura de dados vetor usada na cartografia e no SIG. Apresenta ainda alguns aspectos relativos a confecção de mapas com CAD: aponta as diferenças básicas entre CAD e SIG; menciona como se procede ao registro dos dados no CAD; como são relacionados às entidades do CAD com as feições do SIG; tece alguns

comentários de como é a estrutura e organização do desenho CAD e como pode ser realizada a conversão de CAD para SIG. Também apresenta um tópico sobre a estruturação de dados vetor em CAD. Por fim apresenta a metodologia de digitalização usada na DSG discutindo detalhes e fases importantes deste tipo de aquisição de dados.

O terceiro tema versa sobre a qualidade de dados para SIG. Aborda em seu corpo os entraves para o uso do SIG a as principais fontes de erros. Aborda ainda mais profundamente os erros oriundos da digitalização vetorial de uma carta passando pelos erros associados: com a fonte da carta; com a representação digital da carta, a sobreposição de camadas contendo elementos de área e a erros comuns ocasionados durante a aquisição de dados. No seu encerramento é apresentando o algoritmo *Douglas-Peucker*, amplamente usado para eliminação de vértices em arquivos de estrutura vetorial.

O quarto tema versa sobre a tecnologia de Banco de Dados Orientado a Objeto (BDOO). Apresenta inicialmente os conceitos básicos relacionados ao assunto, tais como, objetos e identidade, encapsulamento, métodos, tipos e classes, herança, polimorfismo, dentre outros. Aborda ainda alguns aspectos relacionados à modelagem de dados orientada a objetos para fins cartográficos, apresentando a modelagem de dados desenvolvida pela DSG para fins de armazenamento e produção de dados. Tem, entretanto, como foco principal a apresentação e descrição do sistema orientado a objetos *Gothic* o que é realizado por meio de comentários a respeito dos aspectos relacionados à arquitetura do sistema e como o sistema entende e lida com os dados geográficos. Descreve também como o BDOO do *Gothic* aborda os conceitos da tecnologia orientada a objetos, como trata a topologia e a questão de versionamento de dados.

O Capítulo 3 trata da validação de dados em ambiente de CAD. Aborda inicialmente a validação de banco de dados geográficos levantando questões importantes relacionadas ao assunto. Expõe de maneira mais ampla a estruturação e validação de dados vetoriais para SIG adotada pela DSG, dissecando a metodologia empregada de maneira a abordar todos os aspectos considerados importantes, tais como os critérios para criação de classes de elementos espaciais e a definição das categorias estabelecidas na Tabela da Base Cartográfica Digital (TBCD) criada e usada pela DSG. Expõe ainda sinteticamente o fluxo de dados geral do processo produtivo, bem como, o fluxo da fase de validação e

estruturação. Apresenta e descreve as principais rotinas desenvolvidas para execução da fase de validação e estruturação, bem como, todos os procedimentos e seqüências de operações que devem ser executadas.

O Capítulo 4 aborda a questão da validação no ambiente orientado a objetos, o experimento realizado e as conclusões preliminares chegadas. No seu corpo é encontrada uma descrição de como foram desenvolvidas as rotinas computacionais implementadas no ambiente, bem como, da linguagem de programação LULL (*Laser-Scan User Language* – Linguagem do Usuário Laser-Scan) usada para escrever os códigos. Descreve ainda a metodologia de validação criada para a validação no ambiente orientado a objetos juntamente com uma descrição mais completa de todas as rotinas desenvolvidas e estratégias adotadas para o seu desenvolvimento. Apresentam os resultados obtidos pela aplicação da metodologia e rotinas desenvolvidas em situação de validação de dados oriundos da aquisição, as considerações finais e encerra a dissertação com as conclusões encontradas.

O Capítulo 5 apresenta as conclusões do trabalho e as recomendações para futuras abordagens ao assunto discutido, bem como sugestões para possíveis pesquisas relacionadas ao ambiente orientado a objeto usado.

Capítulo 2 – CONCEITUAÇÃO TEÓRICA

Parte I – CARTOGRAFIA AUTOMATIZADA E SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

2.1 Sistemas de informação geográfica e cartografia

A necessidade da humanidade de conhecer a distribuição do espaço no qual está inserida tem sido constante desde os tempos mais remotos. Por meio de tal conhecimento as civilizações têm garantido condições para sua sobrevivência e desenvolvimento ao longo da história.

Concomitantemente, a comparação de dados sobre a distribuição espacial e as propriedades da superfície da Terra tem sido uma das atividades importantes das sociedades organizadas. Desta comparação surgiram duas áreas de conhecimentos distintas, porém muito próximas a Geodésia e a Cartografia. Ambas as áreas tratam de assuntos correlacionados e seu desenvolvimento acompanha o próprio progresso da civilização.

A geodésia trata do estudo da forma e dimensões da Terra, e tem como origem os experimentos desenvolvidos pelo matemático grego Eratosthenes datado do século VI A.C.. Hoje esta ciência se encontra bem evoluída e abrange como sua área de estudo também o campo gravitacional terrestre.

A cartografia por sua vez apareceu no seu estágio mais elementar sob a forma de mapas itinerário feitos pelas populações nômades da antiguidade. Esta passou a ser considerada como ciência com o advento das grandes navegações e conseqüente desenvolvimento das relações comerciais entre os povos.

Durante o século XIX surgiram as primeiras instituições em nível governamental responsáveis pela execução de mapeamento das nações de forma sistemática. Cada país criou seu Serviço Geográfico, o qual era responsável originariamente, pela produção de mapas topográficos de todo o território nacional, atividade esta, desempenhada até hoje. Surgiram assim os primeiros mapas topográficos

padronizados, os quais se tornaram as principais fontes de dados espaciais. As informações apresentadas por esses mapas incluem basicamente os acidentes naturais e artificiais encontrados na superfície terrestre, e têm por finalidade descrever numa determinada escala com a maior exatidão possível todo o espaço mapeado.

Com o advento da informática no final da década de 80 e início da de 90 ocorreu uma revolução na produção de dados espaciais. As técnicas cartográficas de aquisição, tratamento e edição migraram para aquele ambiente permitindo explorar possibilidades de automação de processos anteriormente impossíveis. Com a explosão da era da informação e a popularização dos SIG a procura por tais tipos de dados aumentou significativamente impondo as instituições responsáveis pela confecção de mapas, governamentais ou não, um ritmo de produção jamais imaginado.

Como os SIG utilizam mapas como matéria prima para realizar consultas e análises envolvendo o contexto espacial e estes por sua vez representam graficamente dados espaciais, é importante introduzir aqui alguns conceitos relacionados à ciência que cuida da produção de tais produtos, ou seja, a Cartografia, bem como realizar alguns esclarecimentos a respeito de designações de alguns termos usados para definir tecnologias e áreas do conhecimento relacionadas ao assunto.

2.1.1 Mapas para visualizar o mundo

É difícil conceber o panorama de uma área por qualquer outro meio que não seja pela consulta a um mapa. Um mapa posiciona dados espaciais, ou seja, dados sobre objetos ou fenômenos para os quais a localização sobre a terra é conhecida, em seus corretos relacionamentos uns aos outros. Um mapa pode ser considerado como um sistema de informação que fornece respostas para muitas questões concebidas à área desenhada: a distância entre dois pontos, a posição de pontos em relação a outros, a dimensão das áreas e a natureza dos padrões de distribuição. As respostas podem ser lidas diretamente da imagem do mapa, sem a necessidade de teclados ou a de abrir alguns arquivos.

Teoricamente, os sistemas de informações geográficas não seriam capazes de solucionar diversos problemas sem mapas, pois o que distingue e dá grande versatilidade a estes sistemas é justamente a possibilidade de fornecer as informações em termos espaciais.

O termo ‘mapa’ é usado em muitas áreas da ciência como um sinônimo para um modelo do que ele representa, um modelo o qual possibilita a percepção da estrutura de um fenômeno representado. Assim o mapa é mais que um divulgador, ele também possibilita conhecer o fenômeno está para ser mapeado. Por ‘método cartográfico’ entende-se o método de representar um fenômeno ou uma área de maneira que sua estrutura espacial seja visualizada e esta usualmente permita algum experimento. Quando se representa informação espacial na forma de um mapa o mesmo tem que se limitar, ao espaço disponível, ao essencial e acerca de qual é a estrutura da informação.

2.1.2 Cartografia: Definição, Objeto e Objetivo

O conhecimento da distribuição das feições encontradas no espaço geográfico é de suma importância para as atividades humanas, dele depende o planejamento para a elaboração de projetos destinados à melhoria das condições de existência das sociedades. Dentro deste contexto, a cartografia como ciência desempenha um papel extremamente importante, pois por meio dela é possível obter uma representação do espaço de maneira confiável e adequada.

A confiabilidade está ligada diretamente a assuntos relacionados à exatidão da representação tanto em termos posicionais como em termos de conhecimento dos atributos das feições mapeadas. Já adequação esta ligada ao modo como as feições foram representadas, assunto relacionado aos princípios da *semiologia* gráfica, conceito a ser abordado mais a diante.

Nas bibliografias ligadas ao assunto, dependendo do autor, pode se encontrar diversas definições para o termo ‘Cartografia’. A seguir serão apresentadas algumas adotadas por órgãos de representatividade mundial e que podem ser amplamente empregados.

Num sentido mais amplo a Organização das Nações Unidas (ONU) define cartografia como segue:

“Cartografia é a ciência que tem por objetivo elaborar toda espécie de cartas e plantas, compreendendo o conjunto de operações desde os levantamentos iniciais até a impressão final”. [ONU, 1949]

Já num sentido mais restrito a UNESCO estabelece a definição apresentada abaixo:

"Cartografia é o conjunto de ciências, técnicas e artes que intervêm a partir dos resultados de observações diretas ou da análise de documentos existentes, tendo em vista a elaboração e a preparação de mapas, plantas e outras formas de representação cartográfica, bem como a sua utilização". [UNESCO, 1966]

Ao analisar as duas definições fica explicito o termo "elaborar", o que nos permite deduzir que a cartografia está intimamente ligada ao processo de criação. Neste caso o processo de criação se faz por meio da representação gráfica de informações apresentadas por intermédio de um mapa ou planta, termos também encontrados nas definições.

Entretanto se fosse necessário resumir a definição de cartografia somente em duas palavras poder-se-ia dizer que Cartografia é "comunicação" e "análise" [Anderson, 1982].

Como "comunicação" concentra-se mais sobre a carta cobrindo aspectos da maneira como foi feita e como pode ser lida e interpretada. É uma forma de comunicação especializada que dá ênfase em disponibilizar a informação de forma visual.

Como "análise" está ligada diretamente à cartografia geográfica que concentra no estudo espacial dos fenômenos a serem mapeados. Assim, antecede o mapa, ou utiliza cartas para determinar o conteúdo de outras novas cartas.

O propósito da cartografia, ou seja, o objeto, encontrado nas bibliografias também varia de acordo com a visão adotada pelos diversos autores. Entretanto todos enfatizam o fato da representação gráfica da informação.

De acordo com Saunders (1991), o objeto da cartografia é a representação espacial das combinações e iterações dos fenômenos da natureza e da sociedade, bem como de suas alterações através dos tempos, por um método de informação à base de um sistema de sinais que constituem os símbolos e convenções cartográficas.

Por sua vez Raisz (1969), enfatiza que o objeto da cartografia consiste em reunir e analisar dados e medidas das diversas regiões da terra, e representar graficamente em escala reduzida, os elementos da configuração que possam ser claramente visíveis.

O objetivo da cartografia, portanto pode ser considerado como a visualização da informação espacial, ou seja, a disposição por meio de informação visual da distribuição espacial de elementos importantes ao conhecimento humano.

Este é o fator preponderante que coloca a cartografia à frente de outras ciências quando o assunto abordado é a localização da informação e o contexto geográfico na qual está inserida.

A partir dos pontos de vista apresentados acima é possível descrever a cartografia de acordo com afirmativa de Kraak & Ormeling, (1996) como a comunicação de dados espaciais por meio de mapas. Para ilustrar isto é apresentado um exemplo de modelo de transferência de informação cartográfica na Figura 2.1.

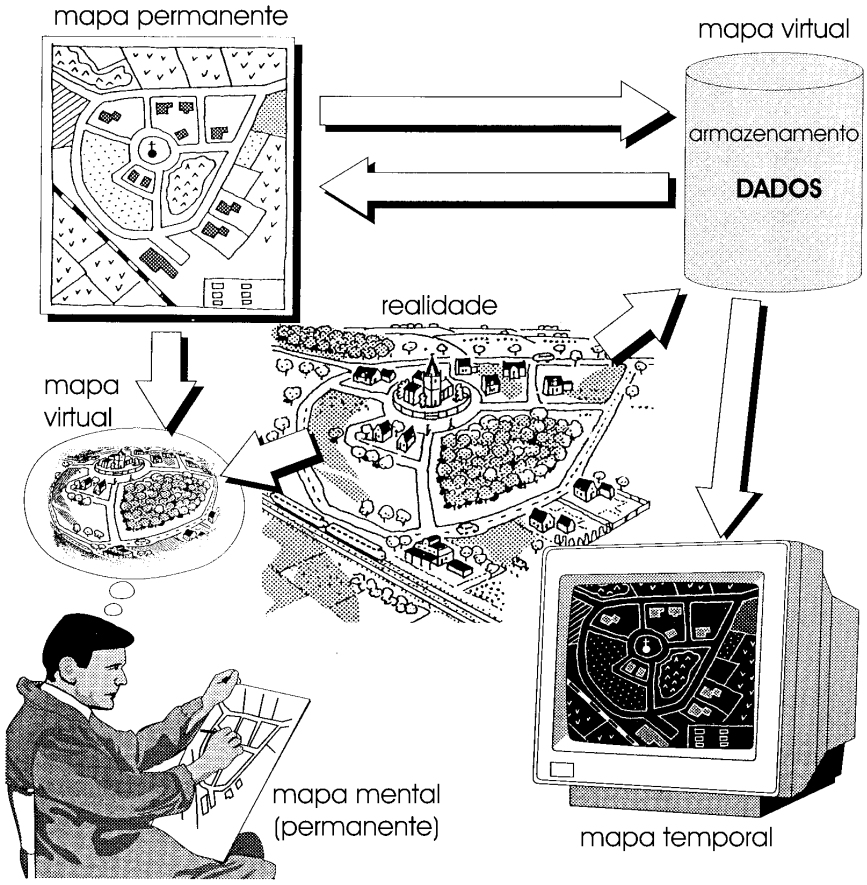


Figura 2.1 - Modelo de transferência de informação cartográfica, com exemplos de mapas permanente, virtual, temporal e mental.
Fonte: Kraak & Ormeling, 1996.

Para finalizar é interessante tecer alguns comentários relacionados aos termos mais genericamente usados para designar os produtos oriundos da cartografia, ou seja: mapa, carta e planta.

Não existe uma diferença rígida entre os conceitos destes termos. É, portanto, difícil estabelecer uma separação definitiva entre o significados dessas designações. Neste trabalho, a distinção entre eles será uma tanto convencional e subordinada à idéia da escala. A seguir segue as definições dos termos apresentados, que serão empregados no decorrer do texto.

De acordo com Saunders (1991), a definição dos termos adotados pelos principais órgãos cartográficos oficiais, a saber, o IBGE e a DSG, são as seguintes:

Mapa – é a representação da Terra em seus aspectos geográficos – naturais e artificiais que se destina a fins culturais ou ilustrativos. São normalmente desenhados em escalas pequenas (menores que 1:1.000.000), abrangendo uma determinada região geográfica.

Carta – é a representação dos aspectos naturais e artificiais da Terra, destinados a fins práticos da atividade humana, permitindo a avaliação precisa de distâncias, direções e altitudes, bem como a localização geográfica de pontos, áreas e detalhes. São desenhadas em escala igual ou maior que 1:1.000.000. Em geral a carta é dividido segundo uma sistemática pré-fixada em folhas, caso em que é denominada “carta normal”.

Planta – Quando a superfície representada é tão pequena que possa ser considerada plana, temos uma projeção no chamado “plano topográfico”, que a rigor, não é uma projeção cartográfica. Neste caso a representação é denominada “planta”. Normalmente as plantas são cadastrais ou topográficas, e sua escala é grande.

2.1.3 Processo de comunicação cartográfica

O objetivo primário dos produtos cartográficos é servir de dispositivo para apresentação de uma situação vigente. Eles apresentam uma visão abstrata de uma porção da Terra com uma ênfase sobre algumas feições pré-selecionadas tais como estradas ou rios.

Quando a maioria dos usuários destes produtos, assim como cartógrafos experientes, recebem a tarefa de confeccionar, por exemplo, um mapa, geralmente é

assumido como enfoque para a apresentação do documento cartográfico a finalidade qual o mesmo se destina. Esta prática norteia as discussões sobre como otimizar a geração de um produto cartográfico.

Como a finalidade é o ponto central sobre o qual gira o processo de criação de um mapa, é necessário que o cartógrafo leve em consideração alguns fatores que condicionam a produção de um produto que cumpra seu objetivo adequadamente.

Há uma discussão considerável hoje no contexto da cartografia moderna sobre as regras da comunicação e visualização. A visualização científica, de modo geral, está pensando num método para incorporar computadores (especialmente gráficos) para transformar dados dentro de modelos visuais que não possam ser vistos ordinariamente [Dent, 1999].

David DiBiase [1990] baseando-se neste pressuposto desenvolveu um modelo gráfico de uso nos quais mapas e outros gráficos podem ser inseridos no ramo da ciência (Figura 2.2). Este modelo básico é relevante, não somente para a ciência, mas também para ser aplicado para as análises geográficas e tomada de decisões realizadas com auxílio de SIG.

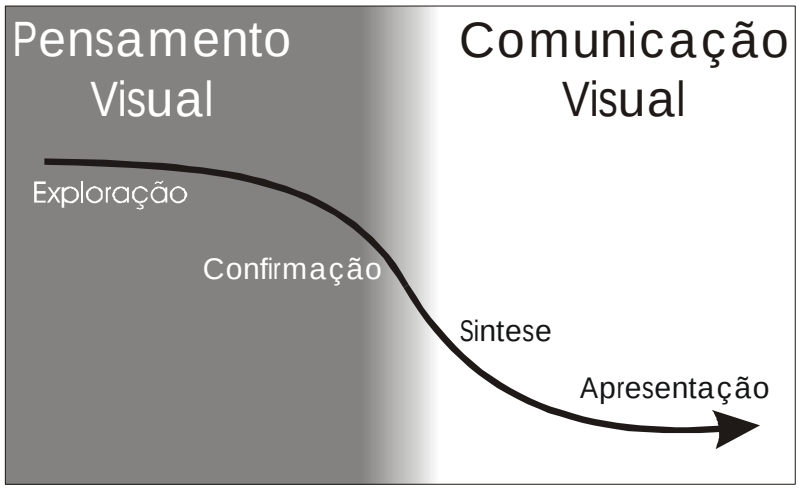


Figura 2.2 - As funções dos gráficos numa seqüência de pesquisa.
Fonte: adaptado de DiBiase,1990.

No modelo de DiBiase, há uma distinção critica entre “pensamento visual” e “comunicação visual”. Mapas podem ser usados para ambos, mas os objetivos da simbolização e desenho diferem.

Para o pensamento visual, os mapas são ferramentas que podem prontamente revelar padrões nos dados e destacar anomalias. O objetivo é ajudar a notificar algo, tal como uma relação entre a localização de uma industria e a incidência de problemas de saúde na população. A decisão da simbolização e desenho feita pelo autor do mapa pode facilitar a habilidade do usuário de notar o desconhecido.

Em contraste a este padrão objetivo de busca, a finalidade da comunicação visual é marcar um ponto, para comunicar o que pensamos que conhecemos. Se uma análise compreensiva usando cartas, dados de saúde, estatística, modelagem, etc. convencer o usuário que há uma ligação entre a distribuição geográfica de um poluente particular e uma espécie de câncer, então se pode criar um mapa para comunicar esta conclusão às autoridades, que baseadas nestas informações podem formular decisões políticas de longo prazo.

O pensamento visual e a comunicação não são objetivos de cartas mutuamente excludentes, mas até certo ponto dois fins de um contínuo ao longo do qual os objetivos mudam gradualmente para diferenças em design e simbolização.

Para entender melhor os principais componentes do modelo gráfico (apresentado na Figura 2.2) os mesmos serão discutidos a seguir a partir do contexto de uma consulta num SIG:

Exploração

Qualquer investigação que envolva o uso de SIG com o objetivo de manipular e examinar dados espaciais pressupõe a existência de mapas em meio digital confiável. Mapas e gráficos neste estágio geralmente não estão documentados porque pertencem ao domínio privado de um pesquisador ou um grupo deles. As representações usadas são genericamente esquemáticas e geralmente não são consideradas de qualidade para publicação. Entretanto esta fase necessariamente precisa ter o contexto de distribuição espacial para a definição do panorama a partir do qual será possível a próxima fase.

Confirmação

Num processo de investigação em andamento, suposições são feitas, questões são levantadas e hipóteses são geradas. Mapas e outras ferramentas de visualização neste

estágio de investigação são geralmente usados como um meio para confirmar suspeitas iniciais. Separadamente ou em combinação com outros gráficos eles podem mostrar o resultado de um esforço de modelagem, ou chamar a atenção para anomalias oriundas da combinação ou transformação das observações originais. Todas estas expressões gráficas orquestram uma função para confirmação ou atribuição de pesos para as hipóteses geradas. Esta é geralmente a visualização de anomalias ou resíduos (casos que permanecem quando uma variável dependente retorna contra uma ou mais variáveis dependentes) que leva ao grande “insight”.

Síntese

Como um analista ganha confiança a respeito de um tema ambiental, os mapas fornecem uma ferramenta pelo meio da qual as idéias podem ser sintetizadas. Eles permitem ao analista produzir uma afirmação coerente, mas abstrata sobre padrões e relacionamentos que não foram levantados. Cada estágio no processamento de dados e de disponibilizá-los representa um processo de abstração. Feições chaves de um nível da análise mais abaixo são salientados e detalhes locais, ruídos randômicos, etc. são filtrados por meio de um esforço para chegar à essência de uma situação. O objetivo da síntese é constituir de detalhes perdidos por meio da recuperação de uma perspectiva (por exemplo, voltando atrás nos dados e especificar o ponto onde o “grande detalhe” tornou-se aparente). A síntese é uma atividade na qual um profissional toma decisões informadas sobre o que enfatizar, o que suprimir e quais os relacionamentos mostrar.

Apresentação

Uma vez convencido da localização (por exemplo, excesso de tóxicos), um atributo numa localização (por exemplo, alcalinidade da superfície da água) ou um relacionamento (por exemplo, entre uma industria e um risco de saúde) e se quer convencer alguém da importância de avaliação realizada, adota-se um mapa por ser um dispositivo de apresentação ideal, para estas situações. Um mapa bem desenhado é importante porque implica em autenticidade. As pessoas acreditam em mapas. Esta aceitação geral pelos mapas (em contraste com a suposição da análise estatística) torna esta ferramenta um dispositivo poderoso, que deve ser construído e usado com cuidado.

A visão de David DiBiase inclui a comunicação na porção do domínio público do modelo. Seu modelo também sugere que a visualização ocupa um lugar ao longo do contínuo, com a exploração e confirmação no domínio privado, e síntese e apresentação no domínio público. Para ele o domínio privado se constitui do pensamento visual e o domínio público da comunicação visual.

De acordo com Alan MacEachren este ponto de vista está usando o termo visualização numa nova maneira para descrever a cartografia como uma ferramenta de pesquisa.

O próprio MacEachren coloca a comunicação cartográfica dentro de um "cubo cartográfico" posicionando-a em um dos seus vértices, enquanto a visualização ocupa o vértice oposto. A Figura 2.3 ilustra o cubo a partir das componentes do modelo de DiBiase no qual a apresentação (encontrada na porção do pensamento visual) aparece próxima ao vértice inferior da figura geométrica e a exploração (encontrada na porção da comunicação visual) é posicionada próxima o vértice oposto. Ainda de acordo com o esquema gráfico apresentado a análise se posiciona entre os dois extremos, o que significa que esta atividade depende da iteração entre as duas porções, fato comprovado em SIG.

Um aspecto interessante na discussão entre cartografia e visualização é que a comunicação cartográfica não está morta, mas incorporada dentro de descrições cartográficas muito mais complexas, certamente, como um componente importante.

Como MacEachren, próprio diz, "Todos os autores, entretanto, parecem estar de acordo que a visualização inclui tanto componentes do pensamento visual/análise como apresentação/comunicação e sugerem (ou ao menos implicam) que a comunicação é um sub componente da visualização" [MacEachren, 1994].

Cabe aqui ressaltar que todo o processo apresentado neste tópico pode sofrer de ruídos e distorções caso os dados armazenados nos bancos de dados dos SIG apresentem inconsistências. Para que este fato não prejudique a tomada de decisão pela indução ao erro é necessário o estabelecimento de controle de qualidade capaz de detectar as inconsistências e providenciar as devidas correções. Para tanto se faz impor um tratamento adequado dos dados com a finalidade de minimizar os erros para que possam ser utilizados nos SIG, particularmente aqueles de natureza espacial.

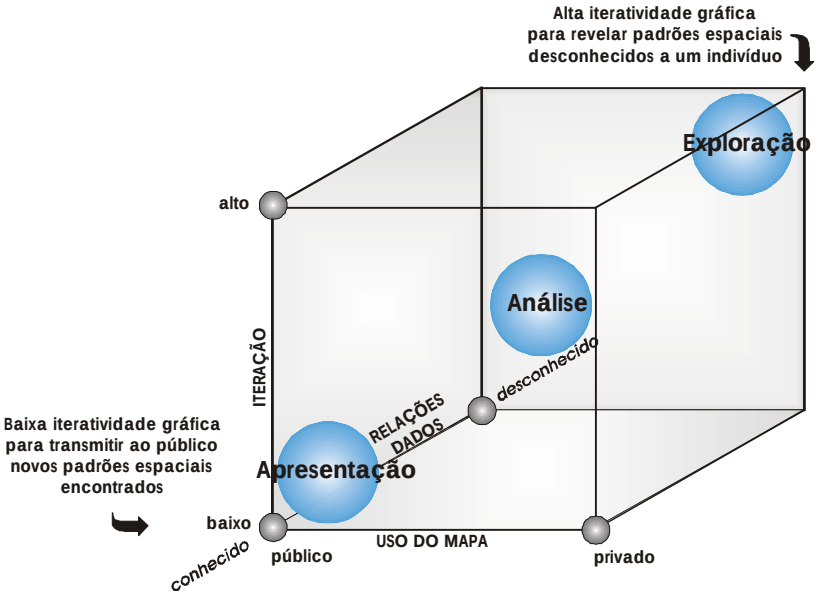


Figura 2.3 - O “cubo cartográfico”
Fonte: adaptado de MacEachren, 1994.

2.1.4 Dados espaciais

Sistemas de informação geográfica são diferentes de outros sistemas de informação, naqueles os dados contidos possuem uma característica especial, que permite referenciar objetos ou fenômenos em relação a uma posição específica no espaço, o que permite a sua localização. Por causa desta característica especial a localização de objetos e fenômenos podem ser visualizados, e estas visualizações – chamadas mapas – são a chave para seu posterior estudo. A Figura 2.4 mostra como os objetos do mundo real que podem ser localizados no espaço (tais como casas, estradas, campos ou montanhas) podem ser abstraídos do mundo real como um modelo de paisagem digital, de acordo com algum critério predeterminado, e armazenado num sistema de informações geográficas (como pontos, linhas, áreas ou volumes) e mais tarde (após ter sido convertido para dentro de um modelo cartográfico digital) representado num mapa (com convenções pré-estabelecidas) e integrar as idéias das pessoas sobre o espaço. Quando armazenados num SIG estes dados espaciais são usualmente divididos em dois tipos de dados, os dados geométricos e dados de atributos. Os primeiros referem aos aspectos geométricos (localização e dimensões) dos fenômenos enquanto que o segundo referem a outros aspectos de características não geométricas.

O conjunto de dados armazenados de uma área de estudo específica é chamado modelo de paisagem digital (MPD). Como este modelo, abstração da realidade, deve ser adequado para comunicação com outras pessoas e tem de ser produzido em forma de produtos visualmente palpáveis, é necessário que o mesmo seja convertido para dentro de um modelo cartográfico digital (MCD), o qual consiste de uma série de instruções para que os dispositivos de impressão e visualização (tal como a tela de computadores) possam ser capazes de reproduzir a informação selecionada em forma de mapas.

Para qualificar os dados como “geométricos” ou “georreferenciados” são necessárias certas informações para identificá-los. Estas podem ter origem num sistema de referencia geográfica ou de projeção cartográfica, números códigos que o relacionam a áreas estatísticas, termos topológicos (por exemplo, A está contido entre B e C), ou termos nominais, como o endereço de uma rua. A natureza espacial dos objetos pode ser expressa em suas formas, por exemplo, a forma de abstração com a qual é representado um objeto do mundo real. Há uma subdivisão básica de forma que consiste em ponto, linha, área e volume, e estas podem ser subdivididas, para aparecerem, alongadas, em formato triangular, irregular ou outras. A maneira como estas formas irão aparecer dependerá da escala (ou resolução) do desenho dos dados espaciais, por exemplo, uma cidade pode ser representada por um ponto numa escala pequena enquanto que numa escala grande poderá ser apresentado como uma área irregular.

A maneira como os objetos ou fenômenos do mundo real são abstraídos como discretos ou contínuos é muito importante para subseqüentes procedimentos de armazenamento e mapeamento. Objetos discretos podem ter todos os seus lados definidos, e as coordenadas destes limites podem ser construídas explicitamente. Estes podem então representar a localização de objetos discerníveis (tais como casas, rios, etc.) ou áreas predeterminadas (limites de estados, municípios, bairros, etc.). As representações contínuas são abstrações daqueles fenômenos que são considerados possuidores de uma mudança gradual em valor. Eles podem ser discerníveis ou mensuráveis (tais como dados de precipitação, ou dados do campo gravitacional terrestre) ou ser baseado sobre modelos (como isolinhas de um MDT).

Para posterior procedimento de visualização é essencial que a natureza dos atributos de informação estejam estabelecidos. Estes atributos podem fazer referencias a

características visíveis (por exemplo, área alagada) e características invisíveis (por exemplo, temperatura). Na definição dos valores para os atributos dos objetos por meio de medição ou categorização chega-se a natureza dos mesmos, que pode ser qualitativa ou quantitativa, informação fundamental para a representação final dos dados.

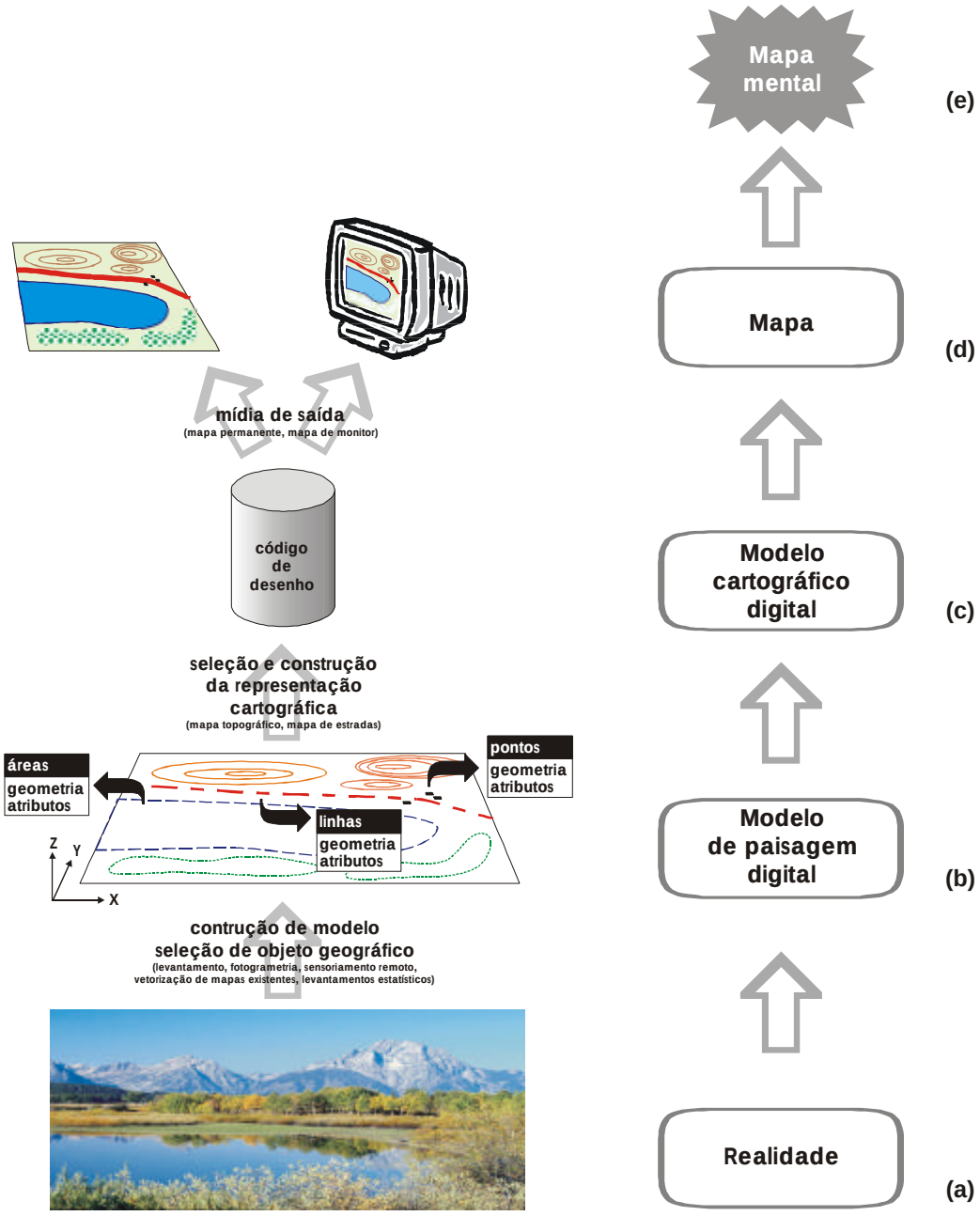


Figura 2.4 - A natureza dos dados espaciais: da realidade (a), via construção de modelo e seleção para um modelo de paisagem digital (b), seguido pela seleção e construção para uma representação cartográfica que dá origem a um modelo cartográfico digital (c), apresentado como um mapa (d), o qual resulta no mapa mental do usuário (e).
Fonte: adaptado de Kraak & Ormeling, 1998.

2.1.5 Sistemas de Informações Geográficas

2.1.5.1 Definição

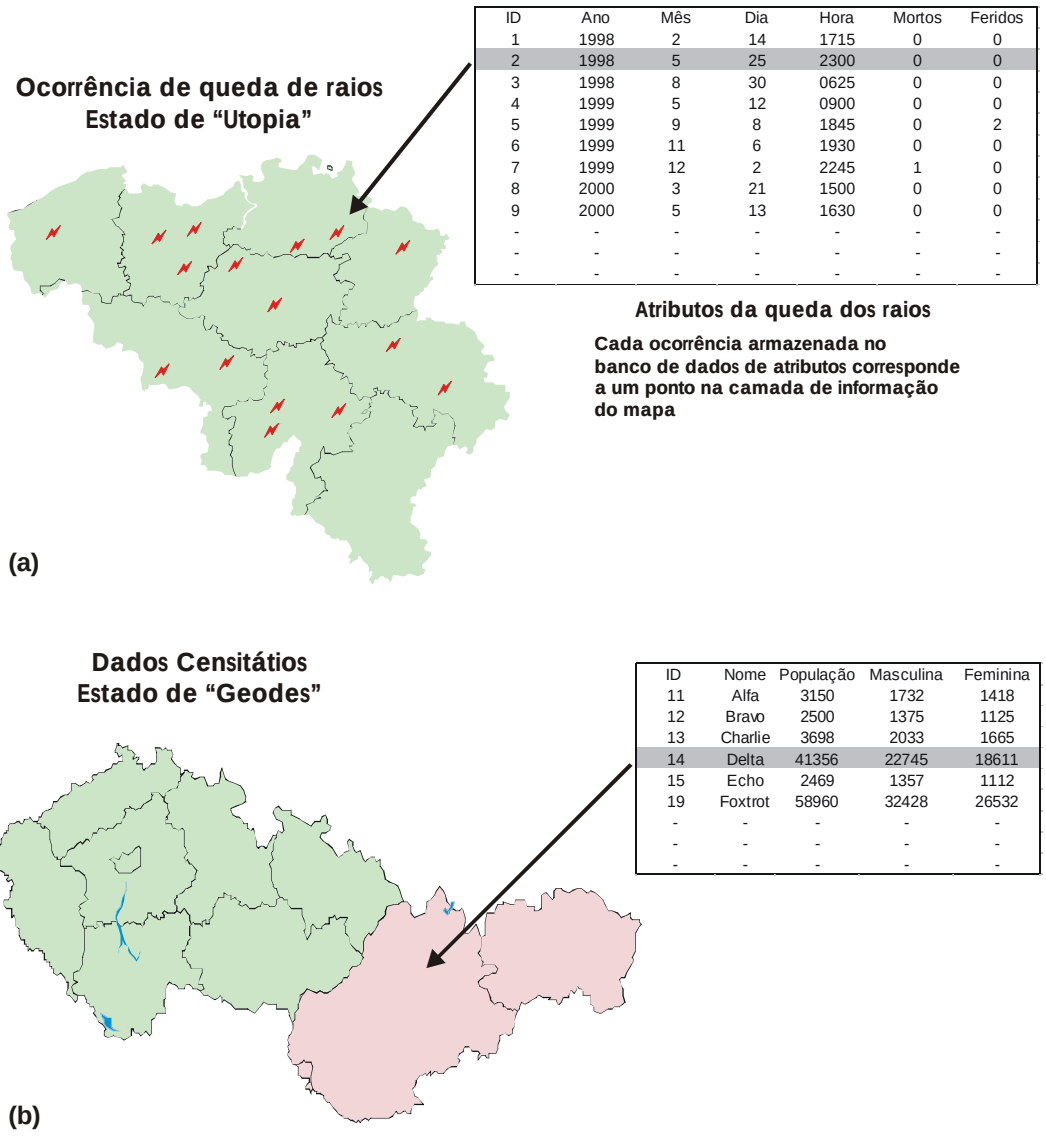
A definição do termo sistema de informações geográficas, ou SIG, tem sido objeto de debate por algum tempo. O SIG tem um grande leque de aplicações e, portanto os usuários usam a tecnologia diferentemente dependendo dos requisitos de suas própria área de estudos particular. Alguns simplificaram a definição de SIG como sendo o uso de pacotes de *softwares.* Outros ampliaram a definição de SIG como sendo a integração de software especializado, hardware e pessoas, designadas para analisar dados geograficamente referenciados ou espaciais. Outros ainda enfatizaram a importância de visualizar o SIG como um processo [Geographer´s Craft].

O processo no qual os dados são adicionados, armazenados, analisados e disponibilizados são requerimentos específicos de cada projeto. O processo SIG é, entretanto, único para cada projeto. Considerando estes pontos de vistas, um sistema de informações geográficas é definido como um processo assistido por computador designado para adicionar, armazenar, analisar e disponibilizar dados espaciais e seus atributos. A Figura 2.5 ilustra o relacionamento dos dados espaciais com tabelas contendo atributos alfanuméricos contidos num banco de dados para duas aplicações distintas: (a) Ocorrência de queda de raios no estado fictício de Utopia, onde cada queda de raio esta cadastrada num banco de dados com atributos específicos e é associada a uma referencia espacial. (b) Informações do senso do estado fictício de Geodes, onde cada município contido numa base cartográfica possui associados os dados de população total, masculina e feminina. A componente chave nesta definição é a habilidade do SIG para analisar dados espaciais. A análise baseada na locação é o que distingue o SIG de outros softwares de mapeamento e gerenciadores de banco de dados [Eastman, 1997].

O SIG tem sido descrito como uma tecnologia poderosa, e seu desenvolvimento comparado à introdução da prensa, do telefone ou o primeiro computador [DeMers, 1997]. O SIG é uma importante ferramenta porque ele permite a manipulação e disponibiliza os dados geográficos de novas maneiras [Abler, 1988] e melhora a eficiência da mais tradicional análise espacial. A Figura 2.6 exemplifica este poder de manipulação de dados

por meio do cruzamento de camadas de informação com estrutura de dados distinta (dados vetor e matricial).

Com um rápido desenvolvimento, o SIG incorporou conhecimentos da geografia, geodésia, cartografia, ciência da computação, matemática, estatística e inúmeras outras disciplinas. Com o crescente aumento no poder dos microcomputadores e o acréscimo da disponibilidade de uso amigável, o SIG tem se tornado acessível a um grande número de usuários.


Ocorrência de queda de raios
Estado de "Utopia"
(a)
Atributos da queda dos raios
Cada ocorrência armazenada no banco de dados de atributos corresponde a um ponto na camada de informação do mapa
Dados Censitátios
Estado de "Geodes"
(b)

ID	Ano	Mês	Dia	Hora	Mortos	Feridos
1	1998	2	14	1715	0	0
2	1998	5	25	2300	0	0
3	1998	8	30	0625	0	0
4	1999	5	12	0900	0	0
5	1999	9	8	1845	0	2
6	1999	11	6	1930	0	0
7	1999	12	2	2245	1	0
8	2000	3	21	1500	0	0
9	2000	5	13	1630	0	0
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-

ID	Nome	População	Masculina	Feminina
11	Alfa	3150	1732	1418
12	Bravo	2500	1375	1125
13	Charlie	3698	2033	1665
14	Delta	41356	22745	18611
15	Echo	2469	1357	1112
19	Foxtrot	58960	32428	26532
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-

Figura 2.5 - Atributos armazenados num SIG.
Fonte adaptado de Dent, 1999.

A tecnologia SIG utiliza um sistema composto por softwares e hardwares que estão submetidos a uma organização de pessoas interligadas para um mesmo fim, que utilizam dados georreferenciados visando a possibilidade de planejar e monitorar questões ligadas ao espaço físico geográfico através dos produtos gerados pelo sistema, que são arquivos digitais contendo Mapas, Gráficos, Tabelas e Relatórios convencionais. [FatorGis, 2004]

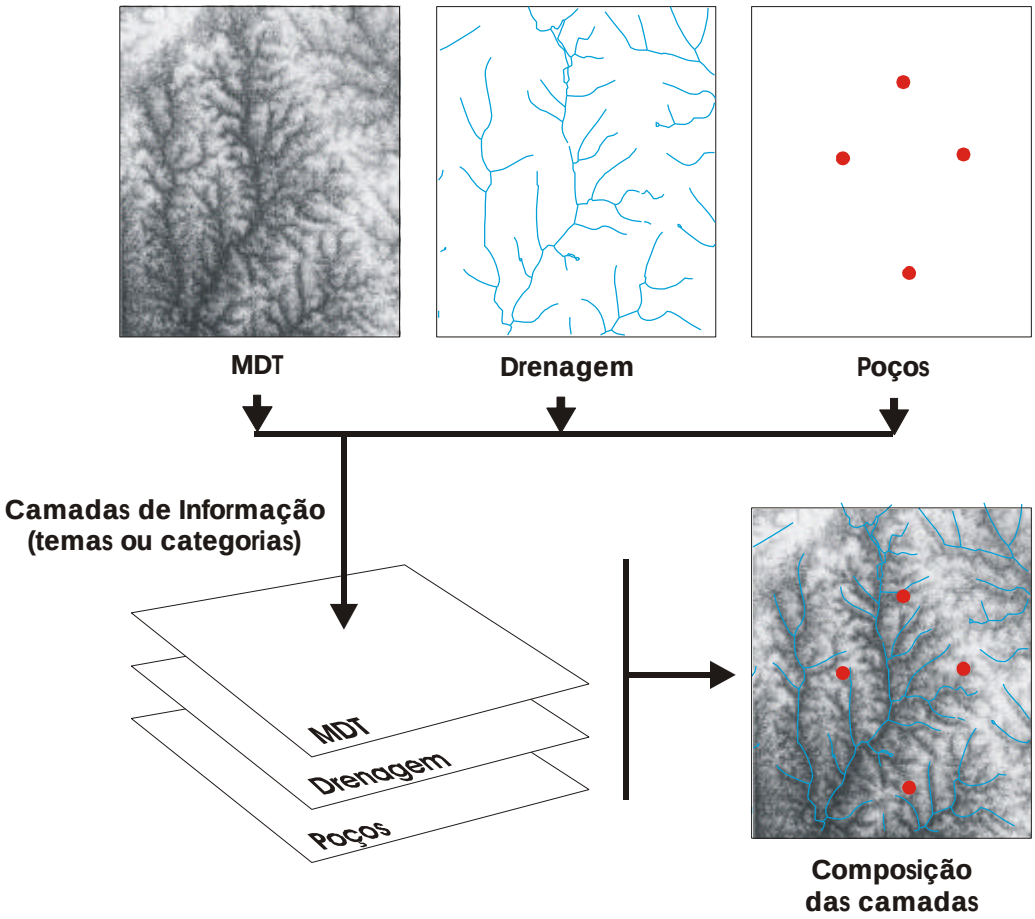


Figura 2.6 - Poder de cruzamento de informação num SIG: sobreposição de dados vetor e raster.
Fonte: adaptado de Kraak & Ormeling, 1998.

O SIG engloba em sua definição vários aspectos já abordados na definição de Geoprocessamento, porém ao GIS, agregam-se ainda os aspectos institucionais, de recursos humanos (*peopleware*) e principalmente a aplicação específica a que se destina.

Portanto é possível definir também o SIG, como um conjunto de ferramentas computacionais compostas de equipamentos e programas que por meio de técnicas, integra dados, pessoas e instituições, de forma a tornar possível a coleta, o armazenamento, o

processamento, a análise e a disponibilização, a partir de dados georreferenciados, de informação produzida por meio das aplicações disponíveis, visando maior facilidade, segurança e agilidade nas atividades humanas referentes ao monitoramento, planejamento e tomada de decisão relativa ao espaço geográfico.[FatorGIS, 2004]. A Figura 2.7 ilustra bem esta definição, que procura mostrar o SIG como uma metodologia integradora de meios, permitindo uma melhor compreensão do que foi mencionado.

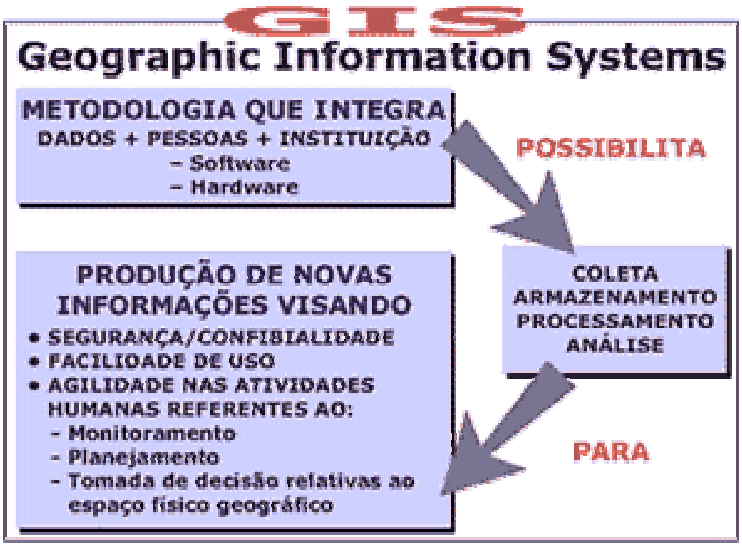


Figura 2.7 - A visão do SIG como metodologia integradora de dados, pessoas e instituições.
Fonte: site da revista FatorGIS, 2005.

2.1.5.2 Componentes de um SIG

Quando as pessoas falam de SIG, geralmente fazem referencia ao software e não à tecnologia. Esta dificuldade de comunicação é percebida com certa freqüência entre os profissionais que utilizam a mesma nomenclatura para se referir a conceitos diferentes. Assim, para um melhor entendimento, se faz necessário explicar informalmente os principais componentes de um SIG, dentre os quais o software é apenas um desses componentes. Para tanto é bastante útil a estrutura apresentada por Burrough que define os componentes como: *software*, *hardware*, dados, usuários e as metodologias de análise (Figura 2.8).

Software – é formado por um conjunto de programas (geridos por um determinado Sistema Operacional), cuja finalidade básica é coletar, armazenar, processar,

analisar e apresentar dados geográficos, tirando partido do aumento da velocidade, facilidade de uso e segurança no manuseio destas informações, apontando para uma perspectiva multi, intra e interdisciplinar de sua utilização. O software contempla basicamente cinco módulos:

1. Coleta, Padronização, Entrada e Validação de Dados;
2. Armazenamento e Recuperação de Dados;
3. Transformação ou Processamento de Dados;
4. Análise e Geração de Informação;
5. Saída e Apresentação de Resultados.

Quanto aos sistemas operacionais, no caso do micro e mini computadores a tendência aponta para o uso das diferentes versões do Windows e do Unix. As estações de trabalho (*workstation*) utilizam sistemas operacionais compatíveis com Unix; os *mainframes* operam sob sistemas VMS, UVS e Unix, entre outros.

Hardware – é o conjunto de equipamentos necessários para que o software possa desempenhar as funções descritas. De forma sucinta, inclui o computador e periféricos, como impressora, *plotter*, *scanner*, unidades de armazenamento (unidades de CD-Rom, disco rígido, fitas magnéticas e ZIP *Drivers*). A comunicação entre computadores também pode ser citada, sendo realizada por meio de um ambiente de rede.

Dados – são matérias primas que alimenta o sistema, permitindo gerar INFORMAÇÃO, que nada mais é do que o significado (atribuído aos dados), do ponto de vista de um determinado usuário. O poder da informação é, sem dúvida, indiscutível. Porém, o que tem revolucionado os processos tradicionais de utilização da informação é a maneira como ela pode ser rapidamente processada e utilizada para diferentes objetivos pelo modo de sua apresentação, ou seja, georreferenciada, ou mapeada.

Usuários – PESSOAS com objetivos comuns formam uma ORGANIZAÇÃO ou GRUPO DE TRABALHO. O SIG por si só não garante a eficiência nem a eficácia de sua aplicação. Como em qualquer organização, ferramentas novas só se tornam eficientes quando se consegue integrá-las adequadamente a todo o processo de trabalho. Para isto não basta apenas investimento, mas o treinamento de pessoal, usuários e dirigentes para maximizar o potencial de uso de uma nova tecnologia.

Metodologias ou Técnicas de Análise – estão diretamente ligadas ao conhecimento e à experiência do profissional que, a partir de um objetivo definido submete seus dados a um tratamento específico, para obter os resultados desejados. Este aspecto mostra que a qualidade dos resultados de um GIS não está ligada somente a sua sofisticação e capacidade de processamento. Muito mais que isso, é proporcional à experiência do usuário.

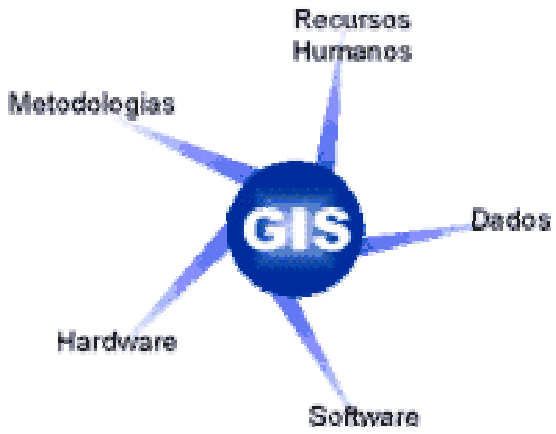


Figura 2.8 - Os principais componentes de um SIG.
Fonte: site da revista Fator GIS, 2005.

Para uma visão mais específica dos componentes de um SIG pode-se adotar a de Dent (1999), que menciona que um típico sistema de informações geográficas é composto de banco de dados espaciais e de atributos para armazenar as informações espaciais e alfanuméricas, hardware e software para entrada de dados, um sistema de análise geográfica e um sistema de visualização de mapa.

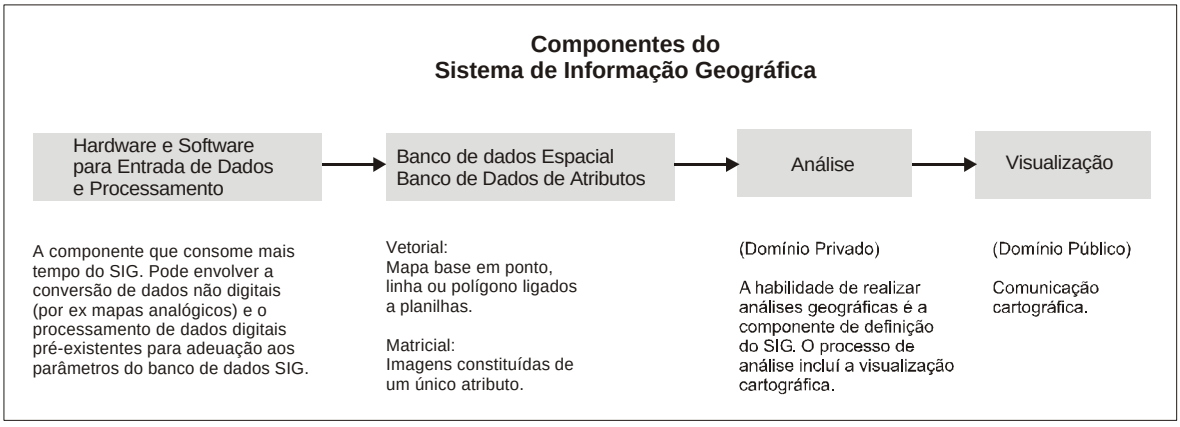

Componentes do
Sistema de Informação Geográfica
Hardware e Software para Entrada de Dados e Processamento
Banco de dados Espacial
Banco de Dados de Atributos
Análise
Visualização
A componente que consome mais tempo do SIG. Pode envolver a conversão de dados não digitais (por ex mapas analógicos) e o processamento de dados digitais pré-existentes para adeuação aos parâmetros do banco de dados SIG.
Vetorial:
Mapa base em ponto, linha ou polígono ligados a planilhas.
Matricial:
Imagens constituídas de um único atributo.
(Domínio Privado)
A habilidade de realizar análises geográficas é a componente de definição do SIG. O processo de análise incluí a visualização cartográfica.
(Domínio Público)
Comunicação cartográfica.

Figura 2.9 - Componentes de um sistema de informações geográficas.
Fonte: adaptado de Dent, 1999.

Dentro desta visão os componentes de um SIG sugerem uma natureza do processo orientado da tecnologia (Figura 2.9). Os dados são introduzidos, manipulados, analisados e disponibilizados de alguma forma. É importante lembrar que o processo, descrito em detalhes a seguir, nem sempre é linear. Análises, por exemplo, podem revelar que são precisos novos dados, necessitando nova introdução de dados e manipulação.

2.1.6 Geotecnologias

Quase sempre ocorre uma confusão no emprego de termos usados para designar o uso de computadores na geração de mapas. A primeira dificuldade já começa com o emprego do termo Cartografia Digital, o qual simplesmente significa que computadores foram de alguma forma usados na geração de mapas. Algumas instituições produtoras de mapas preferem o termo Cartografia Automatizada para designar a mesma operação. Outro termo que designa basicamente a mesma coisa e que surgiu há algum tempo depois do anterior é o termo Cartografia Assistida por Computador.

Como as possibilidades crescem com a produção de mapas em microcomputadores, e com o advento dos desktop publishing, agora já se começa a ouvir o termo desktop mapping (DTM) sendo aplicado às atividades de produção de mapas temáticos de pequena escala. Este termo não deve ser confundido com o DTM (digital terrain model) também usado hoje, no preparo de vistas de mapas tridimensionais. [Dent, 1999]

Como há uma necessidade de padronização de termos visando uma maior clareza será adotada neste trabalho o termo Cartografia Automatizada para designar os trabalhos de produção de dados cartográficos. A adoção deste termo se dá por motivo de praticidade, visto que as principais instituições responsáveis pelo mapeamento sistemático brasileiro o adotam no seu dia a dia.

A Cartografia Automatizada, termo usado amplamente no início das atividades de cartografia em meio digital, surgiu com o advento da informática e sua evolução a qual permitiu a implementação dos processos cartográficos em meio digital. O desenvolvimento desta tecnologia foi impulsionado basicamente por duas comunidades [GIS Course by Digital Earth, 2004]:

- a científica que necessitavam visualizar resultados de modelagens em menor período de tempo, bem como grande quantidade de dados em formato digital; e,
- a cartográfica que vislumbrava a redução do custo e do tempo, de produção e edição de cartas.

Esta tecnologia trouxe benefícios à produção de mapas que anteriormente exigiam um grande dispêndio de tempo e trabalho. As principais vantagens advindas com o surgimento de tal tecnologia, além daquelas já mencionadas, foram basicamente as seguintes:

a. baixo custo de produção, se comparada toda a infra estrutura usada na produção cartográfica convencional;
b. grande facilidade de manuseio do desenho de um mapa, pois mudanças de escala e projeção são mais simples no ambiente digital; e,
c. possibilidade de outros usos para os dados digitais.

Atualmente, a produção cartográfica responsável pela obtenção de dados espaciais lança mão de diversos sistemas. A integração destes sistemas em torno do termo Cartografia Automatizada é bem visto e adotada pelos produtores de dados. Embora o termo se encontre um pouco defasado, devido ao surgimento de diversas tecnologias na área, é fundamental seu resgate devido a sua importância no processo produtivo de dados para os diversos sistemas de informações espaciais. A Figura 2.10 exemplifica a sua importância para os sistemas mais comuns de informações espaciais.

O desenvolvimento de aplicativos para lidar com dados de natureza espacial, tais como o CAD e sistemas de tratamento de imagens de satélite possibilitaram um novo panorama que favoreceu grandemente aos usuários de informações espaciais. Surgiram assim, as hoje conhecidas “geotecnologias”, que amplamente aceitas pela sociedade fez com que o consumo de dados de natureza geográfica aumentasse significativamente.

As “geotecnologias” são o conjunto de tecnologias para coleta, processamento, análise e disponibilização de informação com referência geográfica. Estas são compostas por soluções em hardware, software e peopleware, que juntos se constituem em poderosas ferramentas tanto para a produção cartográfica como para a tomada de decisão.

Dentre elas hoje conhecidas e amplamente empregadas estão o SIG, o Sensoriamento Remoto e o Sistema de Posicionamento por satélites, a Aerofotogrametria, a Geodésia e a Topografia Clássica, e outras mais.

O termo Geoprocessamento surgiu justamente para agrupar numa única tecnologia as diversas "geotecnologias" existentes, mais especificamente aquelas que lidam com informação espacial de natureza geográfica. A adoção deste procedimento permitiu centralizar de maneira formal as mais diversas áreas que trabalham com tais tipos de dados, incluindo aquelas responsáveis pela produção primária de dados espaciais, como é o caso da Cartografia Automatizada.

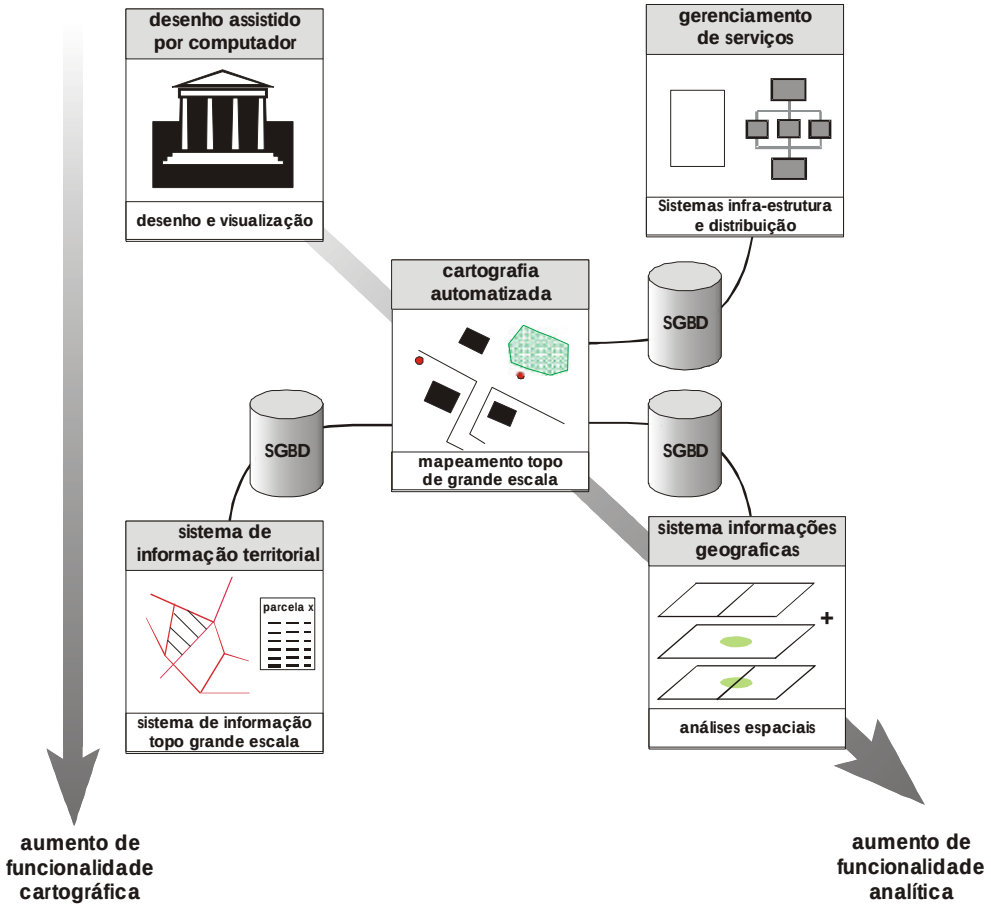


Figura 2.10 - Importância da cartografia automatizada para alguns tipos de sistemas de informação espacial. Fonte: adaptado de Kraak & Ormeling, 1996.

A razão principal da relação interdisciplinar forte entre Cartografia e Geoprocessamento é o espaço geográfico. A Cartografia preocupa-se em apresentar um

modelo de representação de dados para os processos que ocorrem no espaço geográfico. O Geoprocessamento representa a área do conhecimento que utiliza técnicas matemáticas e computacionais, para tratar os processos que ocorrem no espaço geográfico. Isto estabelece de forma clara a relação interdisciplinar entre Cartografia e Geoprocessamento [Câmara, 2002].

Uma razão histórica, que reforça o vínculo mencionado, é a precedência das iniciativas de automação da produção cartográfica em relação aos esforços iniciais de concepção e construção das ferramentas para SIG (veja-se, por exemplo, Maguire et al., 1991). A Figura 2.11 aproveita e sintetiza a discussão apresentada, estendendo-a apropriadamente às áreas de Sensoriamento Remoto, CAD (*Computer Aided Design*) e Gerenciamento de Banco de Dados.

Figura 2.11 - Relações interdisciplinares entre SIG e outras áreas.
Fonte: Maguire et al, 1991.

Como definição simplificada de geoprocessamento pode-se adotar aquela divulgada no site da revista Fator GIS (2005) que o apresenta como:

"O conjunto de técnicas computacionais relacionadas com a coleta, armazenamento e tratamento de informações espaciais ou georreferenciadas, para serem utilizadas em sistemas específicos a cada aplicação que, de alguma forma, utiliza o espaço físico geográfico. Estes sistemas podem ser: GIS, LIS, AM/FM, etc".[Fator GIS, 2005]

Cabe aqui, ressaltar que as informações georreferenciadas neste caso têm como característica principal à localização espacial, ou seja, estão ligadas a uma posição específica do globo terrestre por meio de suas coordenadas geográficas, podendo também ser representadas em termos de coordenadas num sistema de projeção cartográfica qualquer.

Na mesma revista pode-se encontrar ainda uma definição técnica do termo, que também pode ser adotada para um contexto mais específico, o qual o designa como o conjunto de pelo menos quatro categorias de técnicas relacionadas ao tratamento da informação espacial [Fator GIS, 2005]:

- Técnicas para coleta de informação espacial (Sensoriamento Remoto, GPS, Topografia Convencional, Fotogrametria, Levantamento de dados alfanuméricos);
- Técnicas de armazenamento de informação espacial (Bancos de Dados – Orientado a Objetos, Relacional, Hierárquico, etc.);
- Técnicas para tratamento e análise de informação espacial, como Modelagem de Dados, Geoestatística, Aritmética Lógica, Funções topológicas, Redes; e,
- Técnicas para o uso integrado de informação espacial, como os sistemas GIS (*Geographic Information Systems*), LIS (*Land Information Systems*), AM/FM (*Automated Mapping/Facilities Management*), CADD (*Computer-Aided Drafting and Design*).

A idéia expressa por esta última definição pode ser visualizada na Figura 2.12.

Figura 2.12 - Definição técnica de geoprocessamento.
Fonte: site da revista Fator GIS, 2005.

Com o amadurecimento da tecnologia de geoprocessamento surgiram sistemas, que por meio da implementação e desenvolvimento de novas ferramentas, se especializaram na manipulação de dados espaciais visando à realização de análises mais complexas, ao integrar dados de diversas fontes por meio de bancos de dados geográficos. Esses sistemas hoje bem conhecidos receberam a denominação de SIG e são amplamente empregados em trabalhos que exigem a manipulação de dados espaciais de natureza geográfica.

Hoje o Geoprocessamento encontra-se intimamente ligado ao SIG, pois este último engloba praticamente todas as tecnologias empregadas no estabelecimento do primeiro. Embora os dois termos sejam tratados como sinônimos por muitos usuários, existe uma diferença bem clara, quando se adota a ótica dos produtores de dados espaciais. Para esses profissionais o geoprocessamento é o conceito mais abrangente e representa qualquer tipo de processamento de dados georreferenciados, incluindo, os processos e tecnologias relacionadas à aquisição, tratamento e edição de dados gráficos. Já o SIG estaria intimamente ligado aos processos e tecnologias usados no processamento de dados gráficos e não gráficos (alfanuméricos) com ênfase a análises espaciais e modelagens de superfícies.

A adoção deste ponto de vista, entretanto, não deve retirar da mente que os processos envolvidos no geoprocessamento não possam ser empregados numa análise de dados ou que não seja possível o uso do SIG para fins de aquisição, tratamento e edição de dados. O que se procura aqui é fazer uma distinção entre os dois termos para que haja uma melhor compreensão de ambos.

Esta distinção é importante, pois a partir do momento que ela é feita fica fácil entender a maneira com que os diversos sistemas usados na produção cartográfica se comportam e se inter-relacionam no processo produtivo de dados espaciais.

2.1.7 Relação entre SIG e cartografia

Muitos dos conceitos e funções dos SIG foram primeiramente concebidas por cartógrafos [Kraak & Ormeling, 1996]. Estes conceitos englobam desde a maneira como os dados espaciais são apresentados na tela de um computador até o modo como deve ocorrer a entrada dos mesmos. Existe um conflito conceitual entre o relacionamento dos SIG com a

cartografia. Existem duas proposições, a primeira assume que o SIG pertence ao ramo da cartografia que trata da análise de dados visando a produção de cartas. O outro assume que a cartografia faz parte do SIG pertencendo ao ramo que trata da apresentação de dados em forma de mapas.

A distinção entre cartografia, especialmente a cartografia temática, como uma disciplina e um sistema de informação geográfica, é confusa e complicada. Além disso, esta mistura está mudando dramaticamente a fronteira e o alcance das possibilidades na visualização científica.

Há muitas definições de SIG. Previamente, se aceitarmos o ponto de vista, que um "sistema de informação geográfica é um sistema baseado em computador que processa informação geográfica" [DeMers, 1997], verificamos que a função de mapeamento no SIG não fica bem definida. Entretanto, não é incomum hoje em dia encontrar funções completas com várias alternativas para o mapeamento temático em programas de SIG. A facilidade com que dados socioeconômicos podem ser inseridos dentro dos programas também torna muito fácil realizar todos os tipos de mapeamentos temáticos possíveis. Sendo assim, sem entender a conexão entre os dois, o método de processamento de dados, a simbolização e disponibilização em forma de mapas requerem estudos contínuos e exames especializados.
Uma discussão sobre estas duas atividades parece ser a melhor maneira de lançar luz sobre as suas diferenças e similaridades para que a fronteira entre elas possa ser mais bem entendida. O ponto de partida para tal discussão é justamente a dúvida mencionada anteriormente, ou seja:

"A cartografia pode ser considerada parte do SIG ou o contrário, o SIG é que deve ser considerado como parte da cartografia?".

Muitos da comunidade de SIG, especialmente aqueles sem conhecimentos de cartografia, consideram como foco da cartografia as atividades de compilação, *design* e produção, enquanto o SIG enfoca a análise espacial. A partir desta perspectiva, a cartografia e o SIG possuem uma similaridade técnica, mas diferem em objetivos. Por sua vez o conceito tradicional de cartografia, concebido primariamente com o design e produção de mapas, separa o SIG da cartografia. Nesta perspectiva, a cartografia é aplicada somente quando é requerida uma saída gráfica em forma de mapa. De acordo com este conceito, a cartografia pertence a uma pequena parte do SIG.

Embora os mapas impressos tenham sido usados por muito tempo no estudo de padrões espaciais, a ascensão da tecnologia computacional tem possibilitado aos cartógrafos visualizar dados mais eficientemente. Devido ao aumento deste tipo de visualização de dados, modelos mais recentes de cartografia, incluindo aquele deste texto, incorpora análise cartográfica como parte do processo cartográfico. Como enfatizado previamente, a análise de dados espaciais é o foco principal do SIG. Por esta razão, muitos cartógrafos consideram o GIS como parte da cartografia o que vem primariamente de encontro com o pensamento cartográfico de domínio privado. O SIG a partir deste ponto de vista é parte da cartografia.

Entre os novos usuários de mapas, entretanto, o SIG tornou-se um padrão para questionar se realmente há a necessidade de dados espaciais especialmente preparados para aquele sistema. Muitos profissionais, por diversas razões, geralmente dizem que querem introduzi-los num SIG, mas o que realmente significa é que querem um meio de visualizar os dados espaciais, e não uma maneira de analisá-los (objeto fim de um SIG). Neste caso em particular, seria mais interessante para o usuário um conjunto de dados em formato apropriado para a leitura em softwares comerciais comuns e não os dados para SIG, que geralmente possuem um custo maior. O porque da existência desta confusão entre as duas finalidades é uma coisa ainda mal definida.

O fato de os SIG possuírem ferramentas de mapeamento temático tem provocado uma verdadeira euforia entre os usuários destes sistemas que às vezes, não possuindo o devido conhecimento de design de mapas, ocorrem em erros não justificáveis. Isto sem comentar os erros de natureza geométrica cometidos em mapas oriundos de uma análise na qual foram usados dados espaciais de origem duvidosas e não adequadas aos SIG.

Para a proposta deste trabalho de dissertação, portanto, é de extrema importância a perfeita compreensão ‘dos métodos cartográficos’, pois a cartografia estará intimamente ligada a todos os aspectos relacionados ao manuseio de informações geográficas pelas seguintes razões [adaptado Kraak & Ormeling, 1996]:

- os mapas são uma interface direta e iterativa para os SIG;
- os mapas podem ser usados para indexação de fenômenos ou objetos contidos nos SIG;

- os mapas, como meio de visualização, podem ajudar: na exploração visual de dados (possibilitando também a descoberta de padrões e correlações) e na comunicação visual dos resultados oriundos da exploração dos dados nos SIG; e,
- na fase de saída ou disponibilização de resultados, as ferramentas disponíveis nos softwares de cartografia são superiores as funções de saída dos SIG comuns.

Parte II - AQUISIÇÃO DE DADOS GEOGRÁFICOS ARMAZENADOS EM ESTRUTURA VETORIAL

2.2 Aquisição de dados

A aquisição de dados é o ponto inicial para a obtenção de informações geográficas de natureza espacial. Para ambientes baseado em dados digitais desta natureza é fundamental entender como se dá o processo de aquisição. Existem dois métodos básicos de aquisição de dados vetoriais: a digitalização vetorial (também conhecida por vetorização) e a restituição fotogramétrica. A seguir serão abordados alguns aspectos relativos a aquisição deste tipo de dados bem como será descrito de forma minuciosa o método de digitalização vetorial, por ser considerada uma operação comum aos ambientes de SIG.

2.2.1 Entrada de dados no SIG

A entrada de dados é a operação de codificação e escrita dos dados no BD. A manutenção de banco de dados isentos de inconsistências é uma tarefa importante e complexa da qual depende a finalidade do SIG. Dois aspectos dos dados precisam ser considerados separadamente para os sistemas de informações geográficas, estes são primeiro o posicional ou dado geográfico necessário para definir onde o gráfico ou feição cartográfica ocorre, e o segundo, os atributos que grava e o que a feição cartográfica representa. É esta habilidade de processar feições cartográficas em termos de seus atributos espaciais e não espaciais que é a principal distinção entre cartografia automatizada (onde as informações não espaciais relacionam-se principalmente a cor, tipo de linha, símbolos, etc.) e processamento de informações geográficas (onde os dados não espaciais podem gravar usos do solo, características geológicas, proprietários, tipos de vegetação e tantos outros).

A entrada de dados para um sistema de informações geográficas pode ser mais bem descrita sobre três aspectos:

a) entrada de dado espacial (digitalização);

b) entrada de dado não espacial, atributos associados; e,

c) ligação do dado espacial ao não espacial.

Em cada estágio será necessária uma verificação apropriada dos dados por meio de procedimentos de checagem para detectar se o banco de dados está livre de possíveis erros. Estas verificações, dependendo da origem dos dados, podem acarretar prejuízos em tempo e trabalhos de correção complexos que em muitas vezes inviabilizam o uso das bases cartográficas adotadas. A produção de dados espaciais visando seu uso em ambientes de SIG. Portanto deve ser executada prevendo minimização de erros. Para tanto é necessário que o processo de produção seja reestruturado e as fases de produção sejam adaptadas para esta nova realidade.

2.2.2 Estrutura de dado usada na cartografia digital

A cartografia digital tem uma visão do mundo real discriminada na forma de camadas, na qual cada camada representa um aspecto da realidade [Burrogh, 1991], vide Figura 2.13.

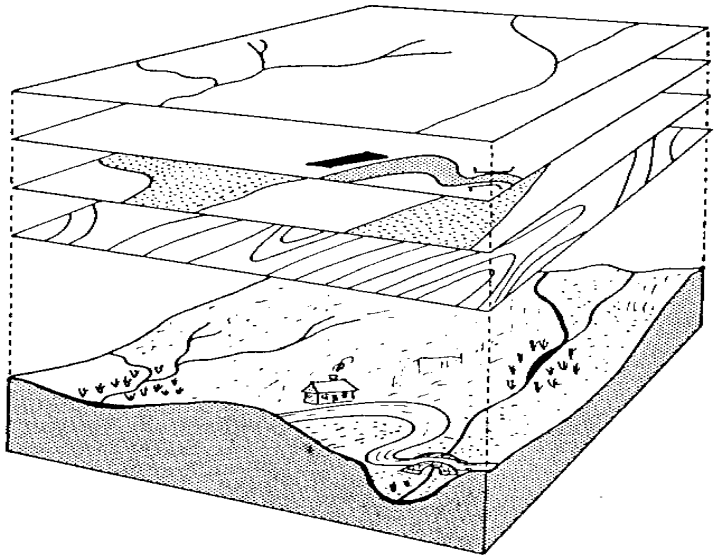


Figura 2.13- O conceito de camadas ou *overlay* do mundo real.
Fonte: Burrough, 1991.

Portanto, as feições do mundo real de caráter geográfico - conhecidas por isso como feições geográficas -, para fins de representação na cartografia digital, são individualizadas e armazenadas separadamente em níveis lógicos. De acordo, a sua natureza espacial, elas serão ainda implantadas através de três modos distintos, a saber: pontual, linear e zonal (Figura 2.14). Portanto, serão estes elementos em conjunto com seus atributos (componente não espacial da informação

geográfica) que deverão ser traduzidos ou codificados para uma estrutura de dados usada pela cartografia em ambiente digital.

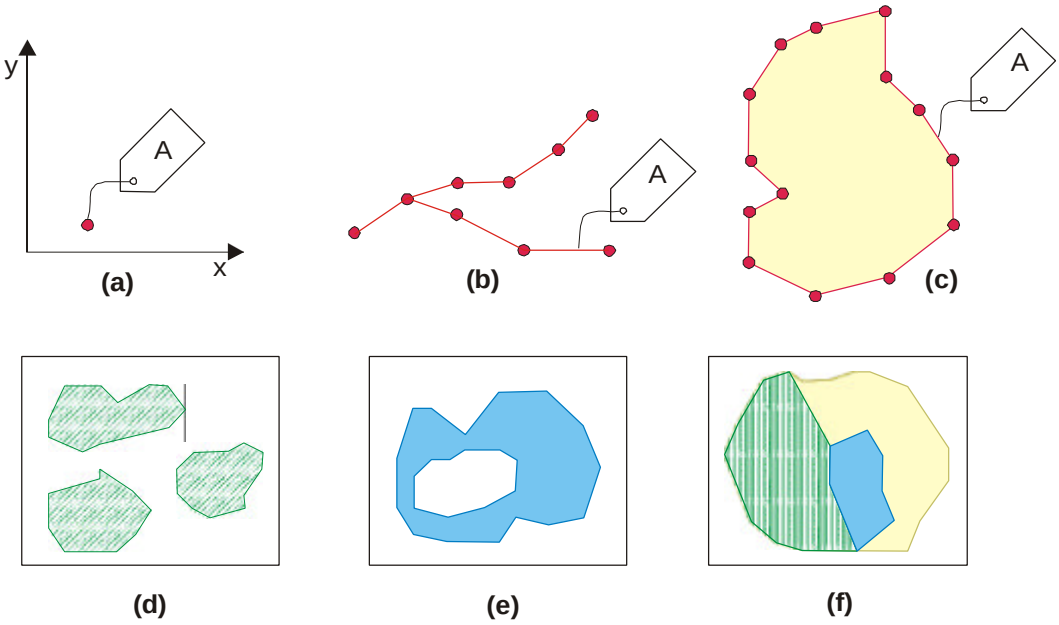


Figura 2.14 - Representação de feições geográficas no ambiente digital: ponto (a), linha (b) e polígono (c). Representação de feições de área e situações mais comuns: áreas disjuntas (d), áreas sobrepostas (e) e as duas situações ao mesmo tempo (f).
Fonte: adaptado de Burrough, 1991.

O mapa digital consiste em uma base de dados computacional finita e discreta, tomando-se necessário converter a realidade geográfica complexa em número finito de registros de acordo com um modelo de estrutura de dados [Goodchild, 1993]. Quanto aos modelos de estrutura de dados, os dois tipos básicos equivalem à representação matricial ou *raster* e a representação vetorial (Figura 2.15).

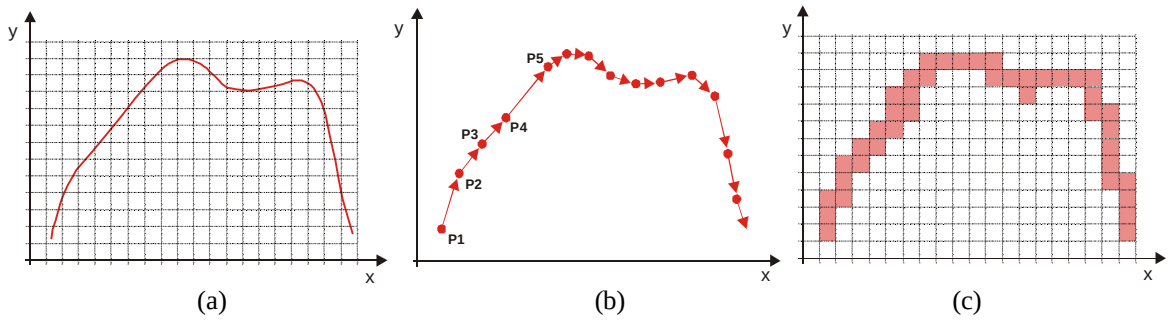


Figura 2.15 Representação de uma linha (a) em estrutura vetorial (b) e em estrutura raster (c).
Fonte: adaptado de Kraak & Ormeling, 1998.

A representação vetorial representa um conjunto georreferenciado de coordenadas que descrevem a localização do objeto no espaço geográfico. Ou seja, um elemento único e distinto pode ser representado por um conjunto de dados

geométricos em um mapa (coordenadas e informação topológica) e pelos seus atributos descritivos, os quais são informações referentes às propriedades não espaciais. Esta representação vetorial consiste então em uma coleção de segmentos de linhas que identificam os limites de pontos, linhas e áreas.

Já, na abordagem *raster*, a informação espacial é traduzida por um arranjo matricial bidimensional, onde cada célula corresponde a uma unidade elementar do espaço geográfico, recebendo um rótulo ou valor. As células possuem uma dimensão específica e cada uma representa uma localização singular. Conjuntos de células de igual valor ou atributo representam a expressão no plano cartográfico de feições pontuais, lineares ou de áreas.

Em suma, a representação *raster* ou matricial pode ser considerada como uma representação explícita, posto que nesta um conjunto de células deve ser sempre explicitado para se definir a localização, forma e área de uma entidade geográfica, enquanto que a vetorial armazena informações apenas sobre os vértices dos limites externos de um objeto, implicando na sua forma e área interior.

O conhecimento dos tipos de modelos de estruturas de dados usados na cartografia digital é fundamental não só para a aquisição dos dados como também para a definição de como os mesmos podem ser armazenados em um SIG. A seguir será abordado com mais detalhe o modelo de estrutura de dados vetorial, por se tratar o foco desta dissertação.

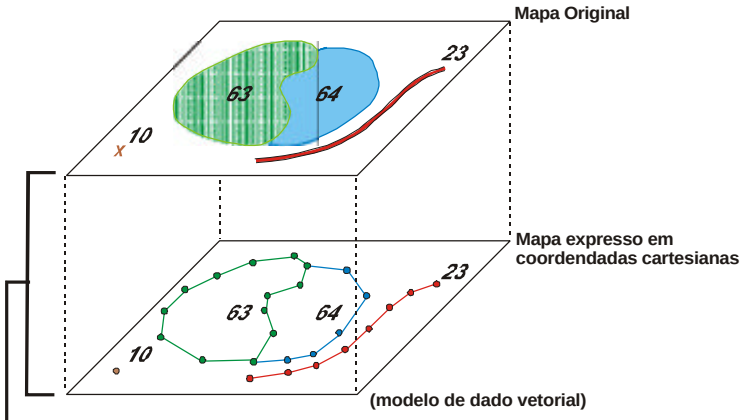
2.2.3 Modelo de estrutura de dados vetorial

O modelo de representação vetorial tem como primitiva principal o **Ponto,** porém, são utilizados três construtores básicos: *o ponto, a linha* e o *polígono. As* coordenadas x e y de um ponto correspondem à localização, em um sistema específico de coordenada, de entidades que são representadas sem dimensões espaciais. A **linha,** formada por uma cadeia de segmentos de linha reta, ou mais especificamente, por uma lista de coordenadas de pontos, é o objeto espacial usado para representar no banco de dados, as entidades da realidade que possuem extensão lineares. O **polígono é** o objeto espacial que representa as entidades com extensões bidimensionais (área), através da definição do contorno da área da entidade. O polígono é formado por uma cadeia fechada de segmentos de linha, podendo ou não, ter outros polígonos embutidos em seu interior.

Existe uma enorme variedade de técnicas de armazenamento de dados espaciais, que são baseadas no modelo vetorial [Laurini, 1992]. Essas técnicas podem ser classificadas de acordo com o tipo de objeto armazenado, ou seja, ponto, linha ou polígono. Características do tipo de aplicações a que se destinam também são importantes. Por exemplo, em um sistema de roteamento de veículos, uma característica fundamental é a conectividade entre as arestas da rede viária, para possibilitar operações de análise de melhor caminho.

Outra classificação existente leva em consideração se os relacionamentos topológicos são, ou não, armazenados. Aronoff (1989), divide os diversos modelos vetoriais em dois grupos: Modelos de Dados *Spaghetti* e Modelos de Dados Topológicos.

Os Modelos de Dados *Spaghetti* utilizam estruturas de dados que armazenam os polígonos/linhas como seqüências de coordenadas de pontos. Nestes modelos, os limites entre duas áreas adjacentes são registrados (digitalizados) e armazenados duas vezes, uma para cada polígono. Estes modelos são utilizados em muitos pacotes de cartografia automatizada onde as informações sobre os relacionamentos entre as entidades não são importantes [NCGIA,1990]. A Figura 2.16 ilustra este tipo de modelo de dados.



Estrutura de Dados

Feição	Número	Localização
Ponto	10	X,Y (Ponto Singular)
Linha	23	X_1Y_1, X_2Y_2, ..., X_nY_n (String)
Polígono	63	X_1Y_1, X_2Y_2, ..., X_1Y_1 (Closed Long)
	64	X_1Y_1, X_2Y_2, ..., X_1Y_1 (Data Structure)

Figura 2.16 - Modelo "espaguete" de dados.
Fonte: adaptado de Pequet, in Taylor, 1991.

A maioria dos SIG utiliza os **Modelos de Dados Topológicos,** os quais usam estruturas de dados que possibilitam o armazenamento de alguns tipos de relacionamentos sendo que a ênfase principal é dada nos relacionamentos de conectividade entre linhas de uma rede (contendo arestas interligadas por nós) e nos relacionamentos de vizinhança entre áreas (representadas por polígonos) adjacentes. Outros tipos de relacionamentos entre objetos espaciais, como por exemplo, "se uma linha cruza uma área" ou se um ponto "está dentro" de uma área, são calculados a partir das coordenadas desses objetos.

A seguir são descritas algumas estruturas de dados vetoriais empregadas nos SIG, de acordo com os tipos de objetos armazenados.

Estrutura de dados para armazenar Pontos

A princípio, as coordenadas (x,y) de posicionamento das entidades com representação pontual podem ser adicionadas como dois atributos extras na tabela de atributos (não-espaciais) que descrevem as entidades. A Figura 2.17 exemplifica esta alternativa. Porém, um objeto espacial do tipo ponto pode ter diversos outros atributos associados a sua representação gráfica, para impressão/exibição em dispositivos de saída. Pode-se citar, por exemplo, o tipo de símbolo que deve ser exibido, a fonte dos caracteres alfanuméricos, o tamanho e a orientação do texto que pode ser exibido próximo ao símbolo, etc.

Uma alternativa também utilizada é manter as informações espaciais em uma tabela e utilizar identificadores de objetos para recuperar os demais atributos não espaciais em uma tabela do banco de dados textual [Burrough, 1986].

Estrutura de dados para armazenar Linhas

As estruturas de armazenamento que visam manter os relacionamentos entre objetos lineares são direcionadas a solucionar problemas em áreas de aplicações que são baseadas em estruturas de rede, como por exemplo, redes de transporte, redes hidrográficas, de distribuição de produtos, redes de infra-estrutura, etc.

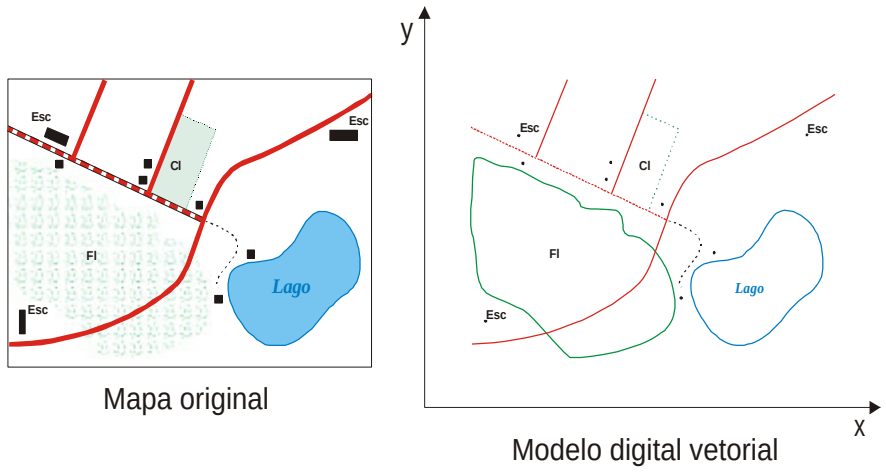
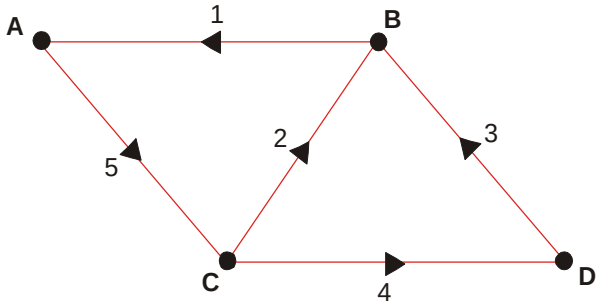


Tabela de Atributos de Pontos

ID	Coord. X	Coord. Y	Nome Escola	Diretora	Fundação	Nr alunos
1	4673000	252500	E.E.Sto Antônio	Jose Rosa	10/03/78	240
2	4674000	253500	E.E.Pe. Anchieta	Maria Lima	07/08/64	765
3	4671000	254000	Instituto de Educação	Rita Reis	03/05/90	360

Figura 2.17 - Armazenamento de dados de pontos: exemplo para feição calssificada como escola.

As redes consistem de dois tipos de objetos espaciais: linhas (ligações, arestas ou arcos) e pontos (nós, interseções ou junções) [NCGIA, 1990]. A Figura 2.18 mostra uma rede composta de 4 nós e 5 arcos (a) e uma estrutura de dados simples (b), capaz de possibilitar a navegação entre os diversos nós da rede.



Arco	Nó de Origem	Nó de Destino
1	B	A
2	C	B
3	D	B
4	C	D
5	A	C

(a)

Nó	Arcos
A	1 -5
B	3 2 -1
C	-4 -2 5
D	4 -3

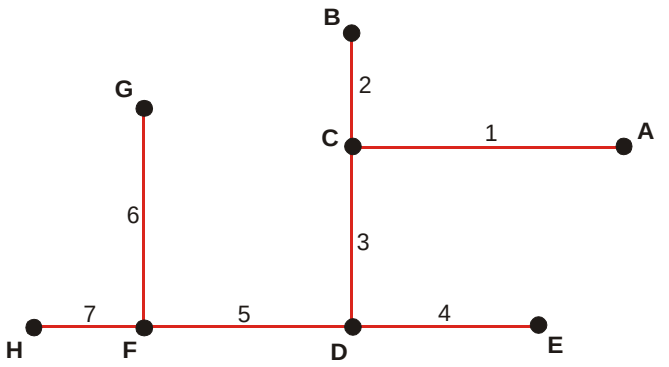
(b)

Figura 2.18 - Estrutura de dados para rede.
Fonte: adaptado de NCGIA, 1990.

Para melhorar a eficiência do algoritmo de navegação, pode ser acrescentada uma nova tabela (c), contendo, para cada nó, a relação dos arcos

adjacentes (números positivos se os arcos atingem o nó e negativo para os arco s que partem do nó).

Uma estrutura de dados para o armazenamento de redes representadas por grafos não-direcionados é mostrada na Figura 2.19. Neste exemplo, descrito em [Laurini, 1992], são empregadas três tabelas, que contém informações sobre os relacionamentos entre arco-arcos, arco-nós e nó-arcos, possibilitando a navegação pela rede em qualquer sentido.



Arco	Conecta aos nós
1	A C
2	B C
3	C D
4	E D
5	D F
6	F G
7	F H

Arco	Conecta ao arcos
1	2 3
2	1 3
3	1 2 4 5
4	3 5
5	3 4 6 7
6	5 7
7	5 6

Nó	Arcos Adjuntos
A	1
B	2
C	1 2 3
D	3 4 5
...	
...	
...	

Figura 2.19 - Representação de grafos não-direcionados.
Fonte: adaptado de Laurini, 1990.

Estrutura de Dados para Armazenar Polígonos

O relacionamento de vizinhança entre entidades bidimensionais (áreas) é representado através de estruturas de dados que armazenam informações sobre polígonos adjacentes. A estratégia mais utilizada é baseada no armazenamento de atributos dos arcos, acrescidos de dois apontadores extras, referentes aos polígonos localizados à esquerda e à direita do arco, percorrido no sentido nó-origem-nó-destino.

A Figura 2.20 mostra um exemplo de uma estrutura de dados simples, que é baseada em três tabelas. A primeira contém os atributos dos polígonos, a segunda os atributos dos arcos e a terceira contêm as coordenadas dos pontos que formam a geometria dos arcos. Este método, utilizado no sistema Arclnfo [ESRI, 1991], tem a desvantagem de não possibilitar a representação de entidades compostas

de mais de um polígono, como por exemplo, um arquipélago, que é uma entidade que precisa ser representada como um conjunto de polígonos.

Este tipo de estrutura possibilita a execução de operações de consulta de maneira bastante eficiente, por não necessitar realizar operações com base nas coordenadas dos objetos. Por exemplo, todos os polígonos, adjacentes ao polígono B podem ser encontrados a partir de consultas à Tabela de Atributos dos Arcos. Cada par de apontadores (polígono à direita, polígono à esquerda), contendo o polígono B indica um polígono adjacente a ele, por ter um arco em comum. Os arcos 3, 4 e 5 identificam, os polígonos adjacentes ao polígono B, sendo que o tratamento dado à área externa aos polígonos (assinalada na figura como polígono D) varia de acordo com a implementação de cada sistema.

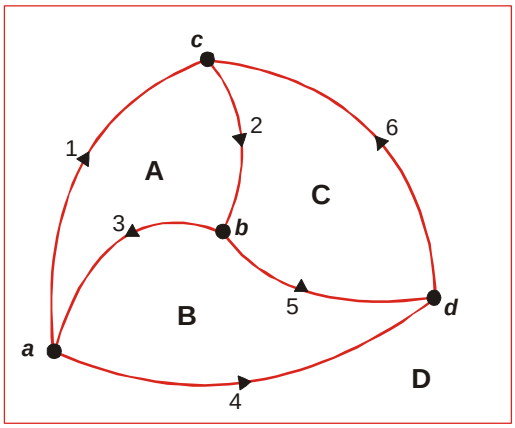


Tabela Atributo de Polígonos

ID	Atrib.1	Atrib.2	Atrib.3	...
A				
B				
C				
D				

Tabela de Atributo de Arcos

Arco	Nó de Origem	Nó de Destino	Área a direita	Área a esquerda
1	*a*	*c*	A	D
2	*c*	*b*	A	C
3	*b*	*a*	A	B
4	*a*	*d*	D	B
5	*b*	*d*	B	C
6	*d*	*c*	D	C

Tabela Geometria de Arcos

Número	Coordenadas
1	$X_1Y_1, X_2Y_2, ...$
2	$X_1Y_1, X_2Y_2, ...$
3	$X_1Y_1, X_2Y_2, ...$
etc.	

Figura 2.20 -Relacionamento de polígonos adjacentes.
Fonte: adaptado de NCGIA, 1990.

Existem muitas variações de estruturas de dados desenvolvidas para o armazenamento da topologia de objetos espaciais. Um exemplo de estrutura mais elaborada é a utilizada pelo sistema CanSIS (*Canadian Soil Information System),* desenvolvida pelo Departamento de Agricultura do Canadá, que utiliza entre outras coisas, uma estrutura para armazenar informações sobre objetos. Um objeto pode conter uma lista de polígonos associada a um único conjunto de atributos, permitindo, por exemplo, um tratamento adequado à representação de um

arquipélago [NCGIA, 19901.

2.2.4 A confecção de mapas com CAD

O CAD *(Computer-Aided Drafting,* ou desenho auxiliado por computador) e o CAM *(Computer-Aided Mapping,* mapeamento auxiliado por computador) são softwares elaborados para criar eficientemente e com muita exatidão as entidades gráficas necessárias para a produção de mapas. Neste meio eles podem ser confundidos com SIG, mas há várias distinções importantes entre CAD e SIG. Embora os programas de CAD não tenham capacidade de executar os tipos de análises realizadas nos SIG, os mapas gerados naquele ambiente provêm à base espacial necessária para que os SIG realizem tais análises. O CAD também pode ser usado efetivamente para editar dados gráficos, geralmente com maior eficiência do que os programas de SIG onde a habilidade de manipular entidades gráficas é secundária à análise.

Com a grande quantidade de dados de mapas que estão sendo produzidos em programas de CAD, é importante para os usuários técnicos de SIG saber como o CAD pode ser usado e a maneira como os dados gerados nestes *softwares* podem ser convertidos para o uso no SIG.

2.2.4.1 Diferenças entre CAD e SIG

Antes de entender como o CAD pode ser usado com SIG, é necessário entender a diferença entre CAD e SIG. Nesta discussão se fará referência ao AutoCAD, produzido pela AutoDesk, um dos pacotes de CAD mais comum, e os pacotes de SIG da ESRI, ArcView, ArcCAD e ARC/INFO.

Embora, a maioria de funções possa ser encontrada em todos os pacotes de CAD e de SIG geralmente usados. A diferença fundamental entre os dois sistemas que deve se ter em mente é que o CAD é projetado para criar e editar entidades gráficas, enquanto o SIG é um banco de dados espaciais que usa gráficos para indicar resultados de análises. A seguir são apresentadas as diferenças básicas entre CAD e SIG:

Criação de topologia

Os softwares de CAD não criam topologias. A topologia é a habilidade

de determinar a conectividade, adjacência e outras relações espaciais entre entidades. Mesmo quando o CAD é usado para conectar linhas *endpoint-to-endpoint* a conexão não é armazenada como a parte de uma rede. Se uma série das linhas for conectada para dar forma a um polígono fechado, o programa de CAD reconhece somente este como um polígono sob circunstâncias espaciais; e não pode reconhecer se um ponto dentro desse polígono está relacionado a este polígono, ou que o polígono tem linhas compartilhadas com polígonos adjacentes.

Sistemas de coordenadas

Enquanto sistemas de CAD são capazes de usar coordenadas georreferenciadas tal como o UTM, muitos desenhos de CAD podem usar preferivelmente coordenadas cartesianas com um ponto de origem arbitrário. Geralmente estes desenhos ou mapas podem ser escalados, movidos, e/ou girado dentro de um sistema de coordenadas georreferenciado. Também, até recentemente muitos sistemas de CAD não possuíam a potencialidade de mudar de sistemas da projeção.

Ligação a atributos

Uma diferença crítica entre os dois ambientes diz respeito à potencialidade de ligar atributo dos dados às entidades gráficas. As entidades do CAD podem ter dados ligados indiretamente, como nomes da camada, cor, largura, comprimento, e em alguns casos o atributo e a informação do texto podem ser ligados às entidades tipo ponto. Entretanto, estes dados não estão prontamente disponíveis em um formato tabular dentro do programa de CAD. Mas um sistema de SIG é projetado para ligar diretamente cada entidade gráfica com um registro em uma tabela dos dados, e em um SIG, a tabela dos dados é mais importante do que o gráfico.

Outras

Há também diversas diferenças menores entre CAD e SIG. Uma delas é que os sistemas de CAD trabalham no nível de exatidão de engenharia (o dobro da precisão), enquanto alguns sistemas de SIG usam um nível de exatidão de mapeamento (precisão única). Os sistemas de CAD também podem usar uma

variedade de tipos de unidade, tais como aquelas destinadas a projetos mecânicos e hidráulicos (pés e polegadas, decimais ou fracionários), enquanto a maioria dos SIG trabalha somente com unidades decimais. Estas diferenças geralmente não causam grandes dificuldades quando o usuário está ciente delas. Uma outra diferença está na estrutura do arquivo; um desenho do CAD está contido geralmente em uma arquivo, mas um "mapa" no SIG está contido geralmente em um ou mais subdiretórios com múltiplos arquivos. Para usuários iniciantes, pode ser dificil compreender que ao trabalhar com um SIG, estão sendo criados múltiplos arquivos de dados e que o que é indicado na tela é justamente uma representação provisória da base de dados do SIG, às vezes extraída das múltiplas fontes de arquivos.

2.2.4.2 O Registro de dados no CAD

Programas de CAD podem ser usados para digitalizar dados de cartas em papel ou sobre imagens destas em tela. A configuração do programa de digitalização em um CAD segue o mesmo principio encontrado em sistemas de SIG para estabelecer marcas de registro. Uma vez que a digitalização é começada em um sistema de CAD, o operador necessita prestar atenção para organização da camada de informação (veja a seção seguinte) e usar ferramentas do CAD tais como *o Snap* para assegurar a topologia correta.

Como os desenhos existentes num CAD não são georreferenciados, freqüentemente é possível trazê-los para um sistema de coordenadas georreferenciadas. Se um desenho for criado em unidades da engenharia (tais como pés e polegadas) o mesmo pode ser transformado para unidades decimais, adotando como ponto de base do redimensionamento às coordenadas 0.0. Se o desenho é produzido em coordenada Cartesiana, e há uma coordenada georreferenciada conhecida para um ponto identificável no mesmo, o mesmo pode ser movido selecionando o ponto identificável como sendo o ponto base e usando as coordenadas georreferenciadas como apontadores de onde o desenho deve se encaixar.

Por exemplo, uma prática comum no CAD é criar um mapa com subdivisões de unidades de engenharia usando as coordenadas cartesianas amarradas a um canto da seção, com as coordenadas arbitrárias de 0.0 (origem das coordenadas). Para registrar este mapa, deve-se aplicar apenas o escalonamento apropriado à escala

das unidades de engenharia, e então mover o desenho inteiro usando o canto da seção como ponto base, para as coordenadas usadas no sistema georreferenciado para aquele canto da seção.

Para desenhos que cobrem uma grande área geográfica, não é possível manter a precisão para todo o desenho ao move-lo tomando por base apenas um ponto conhecido. Enquanto o ponto base usado no movimento se apresenta exato, as entidades nos extremos podem se apresentar fora da posição correta. Se dois pontos forem conhecidos, pode ser possível encontrar resultados aceitáveis usando um comando de rotação.

Se ainda permanecerem distorções, então será necessário aplicar transformações ou técnicas de ajustamento de coordenadas que usam múltiplos pontos. Nem todos os sistemas de CAD possuem esta capacidade, e em alguns casos, os dados precisarão ser levados para um SIG para realizar o registro.

2.2.4.3 Relacionamento das entidades do CAD com as feições do SIG

Há dois pontos fundamentais para atingir o sucesso na conversão de desenhos de CAD para as camadas de informação num SIG. O primeiro diz respeito ao entendimento da relação entre entidades CAD e tipos de feições SIG. Já o segundo diz respeito à maneira como estão tematicamente organizados os dados de CAD. A seguir são discutidos com mais detalhes esses dois pontos.

Entidades CAD x tipos de feições SIG

Em um sistema de CAD existem muitos tipos de entidades, algumas das quais aparecem ser muito similares, mas que podem ser muito diferentes em uso e propriedades. Entretanto, com o SIG, geralmente classificam-se as feições como ponto, linha ou polígonos. O entendimento de como uma entidade no programa de CAD será tratada por um programa de SIG é crítica para o sucesso do uso de dados de CAD num SIG. Uma entidade do CAD que necessita uma explanação adicional é a entidade que representa pontos por símbolos cartográficos. Dependendo do tipo de CAD usado estas entidades recebem nomes diferentes, por exemplo, no AutoCAD é denominado "bloco" enquanto que no MicroStation recebe a denominação de "célula". Um "bloco ou célula" pode ser composto de muitas subentidades que podem ser escaladas durante ou após a inserção numa base cartográfica, apesar desta

característica este tipo de elemento é tratado como uma entidade ponto. Por exemplo, um símbolo de árvore pode ser constituído de poli-linhas com um diâmetro de 10 unidades de medida com atributos gráficos de certa espécie de árvore, podendo ainda possuir inseridas informações a respeito do seu crescimento. Quando o operador está inserindo o símbolo este poderá ser escalado de acordo com o dado relacionado a seu crescimento. Este atributo pode então ir para dentro de um SIG. A Tabela 2.1 ilustra a relação entre feições geográficas, entidades do AutoCAD, conteúdos do Banco de Dados e feições de SIG. Apesar desta tabela fazer referência específica ao AutoCAD e ArcCAD, outros programas de CAD e de SIG poderão apresentar tipos similares de entidade e de feições.

Representação Geográfica	Entidade do AutoCAD	Conteúdos do Banco de Dados	ArcCAD ou Tipo de Feição SIG
Ponto	Ponto Texto Bloco (inserção)	(x,y) coordenadas para locação; Texto pode conter uma linha de atributo de dado; Entidades Bloco podem conter múltiplos atributos; podem conter também informação de escala e ângulo.	*"Point Theme Tic Theme"* para registro das marcas; *"Annotation Theme"* para etiquetar como não dado; Uma tabela de atributos de pontos - *Point Attribute* Table (PAT) - é criada na camada de informação.
Linha	Arc Círculo Linha Poli linha Polígono Sólido 3Dface	Nós-(x,y) coordenadas dos pontos de início e final; ou mudança no valor da linha. Vértices-alterados na direção de uma linha. Isto é similar para os vértices nas poli linhas do AutoCAD.	"Line Theme"; Uma tabela de atributos de arco – Arc *Attribute Table* (AAT) - é criada na camada de informação. ArcCAD criará também um tema de linha como parte de uma camada de informação de linha quando usar DDfeat.
Área-Limite	Arco* Círculo formado por Polígono de linhas Fechado Anel de Poli linhas Fechado Elipse Sólido 3Dface	(x,y) coordenadas das feições do perímetro; Lista de Arcos do Polígono - Polygon Arc List (PAL)	*"Polygon Theme";* Uma tabela de atributos de polígonos - *Polygon Attribute* Table (PAT) - é criada para a camada de informação. A edição das entidades dos limites devem ser feitas para o AAT.
Área-Etiqueta ou Centróide	Ponto Bloco Bloco Atributos Texto	(x,y) do ponto de interseção - usado como posição para visualização dos valores dos atributos	O valor do texto é salvo no "Acad Text Item"; Os atributos são salvos nos "Items" usando o nome de etiqueta de identificação como o nome do item.

Tabela 2.1 Relação entre feições geográficas, entidades do AutoCAD, conteúdos do Banco de Dados e feições de SIG.
Fonte: adaptado de Schaeffer, 2005.

Organização do dado CAD

O CAD usa *layers* (ou camadas) para esta finalidade, outro termo também usado *é levels* (ou níveis). Em ambiente de CAD, as camadas ou níveis são usados na maior parte para controlar a visualização das entidades, definindo cores e tipos de linhas para os diferentes objetos, podendo ativar ou desativar a visualização das camadas ou níveis sobrepondo-as conforme o desejado. Mas uma camada no CAD é também similar a uma camada de SIG, naquela todos os dados da camada são geralmente da mesma classe ou tipo de dados. Por exemplo, todas as entidades de linha usadas para representar estradas podem estar em uma camada chamada "estradas". Ou as linhas poderiam ser mais divididas em classes colocando-as em subcamadas denominadas "Estradas Arteriais", "Estradas Coletoras" ou "Estradas Locais".

Como a maioria dos programas de SIG pode ligar o nome da camada de uma entidade ao registro dessa feição no banco de dados, o nome da camada pode ser usado como um atributo do SIG. Ter um desenho de CAD bem organizado com camadas lógicas nomeadas e classes de feições aumenta grandemente a facilidade e funcionalidade da conversão de dados do CAD para dados de SIG.

2.2.4.4 Conversão de CAD para SIG

De forma geral existem três métodos de usar dados de CAD em programas de SIG. O primeiro é usar programas de SIG que possam ler dados do CAD diretamente, tais como extensões de leitura CAD do ArcView. O segundo método é usar rotinas de tradução usualmente baseadas sobre o formato de arquivos DXF *(Drawing Exchange Format* - formato para troca de dados). O último método é usar um programa que trabalha dentro do CAD para criar a topologia e ligar atributos às feições espaciais. Exemplos deste são o ArcCAD da ESRI, AutoCAD MAP da AutoDesk e GeoMedia da Intergraph.

Leitura direta de dados

ArcView é capaz de ler arquivos do AutoCAD (.DWG e .DXF), bem como arquivos do MicroStation (.DGN). O ArcView leva as entidades desenhadas no CAD para forma de linha, ponto, polígono e anotação das feições, e extrai diversas propriedades de cada entidade tais como cor, nome da camada e tipo de entidade.

Extrairá também os dados de atributos dos símbolos cartográficos (blocos/células). Em alguns casos o ArcView construirá topologias rudimentares para as feições, tal como adjacência, mas não reconhecerá entidades ponto dentro de polígonos como centróides. Uma vez que as entidades de CAD se encontram dentro do *ArcView* é possível fazer operações adicionais tais como, consultas e sobreposições limitadas de camadas. Algumas operações somente serão tratadas após os temas dos desenhos serem convertidos para o formato *ArcView* conhecido como *shapefile (.SHP).* Além disso, após a conversão dos temas para o formato *shapefile* do *ArcView,* os dados podem ser levados para o ARC/INFO usando o comando de *Shapein.*

Tradução a partir do formato DXF

O formato para troca de desenhos DXF *(Drawing Exchange Format* – formato de troca de desenhos) é um formato de arquivo estabelecido pelo AutoCAD e é usado como formato intermediário para migração entre os diferentes programas de CAD. Pelo fato do AutoCAD ser amplamente usado este formato se tomou de fato um padrão, sendo possível sua leitura por diversos pacotes de CAD, SIG e GPS. Como todo tradutor, há ocasionalmente alguma perda da informação entre programas e isto pode, às vezes, causar alguns problemas.

Programas híbridos CAD/SIG

Um programa híbrido CAD/SIG cobre geralmente funções de SIG no programa de CAD, dando aos usuários o melhor dos dois mundos. Por exemplo, o programa ArcCAD da ESRI trabalha com a interface do AutoCAD, a ponto de todas as funções do AutoCAD de criação gráfica, edição e apresentação em tela ser disponibilizadas enquanto possibilita o uso das funções de construção de topologia, pesquisa e análises do ARC/INFO num estilo de menu e caixas de diálogo do AutoCAD. Como o AutoCAD é de fácil customização, possibilitando a criação de diversas ferramentas, isto significa que a maioria das funções do ArcCAD pode ser customizadas para uma melhor produtividade, dos desenhos realizados regularmente no AutoCAD pelos usuários que não possuem o ArcCAD.

As camadas de informação criadas no ArcCAD podem ser exportadas em formato "E00" para uso no ARC/INFO ou outros programas que podem ler os arquivos E00, ou como *Shapefile,* que podem ser lidos diretamente pelo ArcView. Novas

feições criadas como operações do ArcCAD tais como *Buffer*, podem ser visualizadas como entidades do AutoCAD e então usadas.

No geral, um programa híbrido CAD/SIG tal como o ArcCAD é a maneira mais eficiente de lidar com dados CAD dentro de um SIG, desde que você tenha as funções de CAD e SIG disponíveis num mesmo programa. Isto permite ao operador corrigir problemas oriundos de erros do CAD por meio de ferramentas disponíveis no CAD, checar topologias e corrigir problemas de dados de SIG dentro do mesmo ambiente. Os outros dois métodos podem requerer adaptação de metodologia para dividir o trabalho entre os programas de CAD e SIG, para produzir dados para SIG de forma usável, e em alguns casos o usuário pode não obter os resultados requeridos. Mas em muitos casos a transição do CAD para o SIG é completamente viável, especialmente se os desenhos realizados no CAD forem criados corretamente, em sistemas de coordenadas apropriadas, estão bem organizados em camadas e se os relacionamentos entre as entidades de CAD e as feições do SIG dos respectivos programas é perfeitamente entendida.

As feições geográficas existentes em cartas topográficas analógicas antes de serem convertidas para o formato vetorial devem ser estruturadas de acordo com uma definição prévia de atributos gráficos e camadas de informação para serem utilizadas em SIG. A seguir será abordado este procedimento, a partir do contexto de um CAD, por este se tratar, como visto anteriormente, da ferramenta mais utilizada na confecção e transformação do desenho cartográfico em papel para o formato vetorial amplamente usado nos SIG.

2.2.5 Estruturação de dados vetoriais em CAD

As feições discriminadas para serem geocodificadas podem ser agrupadas em categorias ou conjuntos naturais para fins de estruturação dos dados cartográficos por meio do uso da noção de camadas ou *overlay*. Por exemplo, pode-se armazenar os elementos de fundo de mapa em um arquivo digital separado de nome "base" e os elementos temáticos em um arquivo "tema". Isto permite uma maior flexibilidade no manuseio dos dados, como também otimiza o processo operacional de entrada de dados num SIG.

As feições geográficas, muitas vezes simplificadas, como no exemplo da rede de drenagem, são então estruturadas em níveis lógicos, sendo especificada sua

simbolização: estilo de traço, cor, espessura, associação com símbolos previamente criados para preenchimento de áreas, ou representação de linhas e pontos. São também definidos os comandos de digitalização e edição, este último para o caso de poligonização. Para elementos de texto são definidos o tipo de fonte e suas dimensões (horizontal e vertical), em conjunto com o espaçamento e justificação das linhas. Todas as informações mencionadas devem ser relacionadas numa tabela de aquisição de dados a qual tem por finalidade documentar os parâmetros usados na conversão dos dados geográficos.

A tabela de aquisição de dados é definida geralmente com atributos lógicos e geométricos, existentes no ambiente de CAD. Estes atributos são respectivamente:

- ? Nível de informação (também conhecido por camada ou *layer)*, os quais permitiram criar estruturas básicas para as informações de acordo com os temas trabalhados; e,
- ? Atributos gráficos (também denominado atributo de desenho), os quais permitiram individualizar as feições atribuindo um tipo de representação gráfica para cada grupo de elementos.

Os níveis de informação são estruturas básicas encontradas em qualquer CAD, entretanto o número de níveis disponíveis varia conforme *o software* usado. Da mesma maneira, os atributos geométricos encontrados nos tipos de CAD, também podem variar. Dentro dos sistemas de CAD estes atributos são representados por valores numéricos, esse fato permite a individualização das feições, por meio de uma definição prévia que pode ser entendido como uma codificação. Por exemplo, o CAD *MicroStation,* utilizado tanto pelo IBGE como pela DSG, permite individualizar as feições tomando por base os atributos e suas variações conforme a Tabela 2.2.

Atributo	Quantidade	Variação
nível	63	1 a 63
cor	255	0 a 254
estilo	8	1 a 8
peso	16	0 a 15

Tabela 2.2 - Atributos existentes no *software* MicroStation.
Fonte: MicroStation User´s Guide, 2000.

Os aspectos mencionados anteriormente além de permitir a individualização de feições representadas por entes geométricos baseados em linhas e polígonos também são aplicados em feições dadas como pontos. Entretanto a representação gráfica deste último é realizada por meio da aplicação de símbolos previamente construídos e armazenados numa biblioteca.

Esses símbolos são réplicas gráficas em formato vetorial dos símbolos encontrados nas cartas e são construídos numa determinada escala tomada como referencia. Quando empregados para representar entidades pontuais sofrem um redimensionamento para a escala final do mapa por meio de um fator multiplicador.

No âmbito da DSG a biblioteca de símbolos recebe a denominação de biblioteca de células que possuí cerca de 247 tipos de células (pontual e linear) para representar todos os símbolos que podem ser encontrados no espaço geográfico no momento da confecção de uma carta topográfica. A escala de referência adotada na construção desses símbolos é a 1:25.000. Os fatores de dimensionamento adotados para a confecção das cartas do mapeamento sistemático podem ser visualizados na Tabela 2.3 a seguir:

Escala referência	Fator de conversão	Escala do mapeamento sistemático
1:25.000	1,00	1:25.000
	0,50	1:50.000
	0,25	1:100.000
	0,10	1:250.000

Tabela 2.3 - Fatores de conversão adotados para dimensionamento dos simbolos cartográficos contidos na bibilioteca de células.
Fonte: Manual de Vetorização, DSG.

De maneira similar, a biblioteca de fontes, contém as fontes necessárias para reproduzir todos os textos necessários para representar as toponímias encontradas nas cartas. Essas fontes são selecionadas a partir de arquivos de fontes pré-existentes e dependendo do tipo de CAD adotado podem variar.

A Tabela da Base Cartográfica Digital (TBCD) adotada pela DSG, bem como a Mapoteca Topográfica digital (MTD) adotada pelo IBGE, foram criadas tomando por base estas características fundamentais do CAD. A Figura 2.21 ilustra uma parte da TBCD, na qual pode ser visto a definição dos atributos mencionados anteriormente, de acordo com o *CAD MicroStation.*

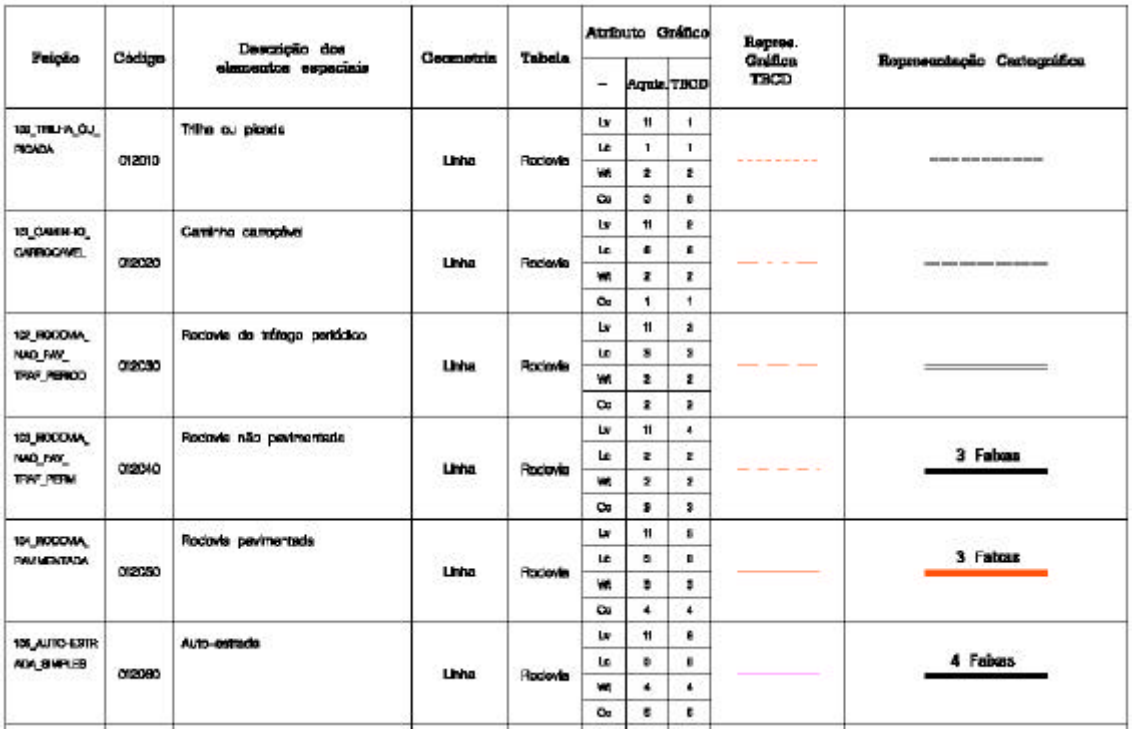


Figura 2.21 Extrato da TBCD indicando os atributos usados
para codificar a feição cartográfica para o modelo digital.
Fonte: DSG, 2002.

A coluna "Aquis." apresentada na Figura 2.21 refere-se a os atributos gráficos da aquisição de dados, Lv *(level* - nível), Lc *(style* - estilo), Wt *(weight* - peso) e Co *(color* - cor). O objetivo final do uso desses atributos é que todas as formas de aquisição de dados sejam, vetorização, atualização por imagem ou restituição, use os mesmos atributos gráficos. Por exemplo, para digitalizar vetorialmente uma rodovia pavimentada seja usados os atributos Lv=11, Lc=0, Wt=3 e Co=4.

A definição dos níveis lógicos, da simbolização e dos comandos associados de digitalização vetorial é definida a priori para um arquivo "semente" (um arquivo contento os ajustes, mas sem dados), o qual é usado para a criação de outros arquivos de desenho cartográfico.

O arquivo "semente" também contém as especificações dos parâmetros de digitalização, da projeção cartográfica, das unidades de trabalho e da resolução cartográfica.

Após as definições necessárias, inicia-se o processo de digitalização e edição gráfica dos elementos cartográficos, obedecendo a uma seqüência

estabelecida em função da prioridade da informação e da mobilidade do elemento gráfico. Desse modo, garante-se a não superposição de símbolos em conjunto com uma maior clareza de representação.

2.2.6 Aquisição de dados vetoriais para SIG via CAD

A aquisição dos dados espaciais em formato vetor no ambiente digital é realizado por processo de digitalização vetorial que pode ser baseado em diversos métodos. É fornecida abaixo uma descrição mais detalhada deste processo, assim como, descrito o método mais usado, a saber, a vetorização.

2.2.6.1 Processo de digitalização vetorial ou vetorização

O processo de digitalização vetorial, conhecido simplesmente como vetorização, é a forma mais comum para obtenção de dados geográficos no formato de vetores para uso em SIG. Para que o processo de vetorização possa ser aplicado sobre documentos cartográficos basta que estes existam, estejam em estado de conservação adequada e em condições de ser submetido ao processo de conversão.

Neste método uma carta depois de convertida para formato matricial (também conhecido por *raster)* é transformada por meio de pontos, que são calculados e interpretados como pares de coordenadas x e y em primitivas geométricas (ponto, linha e polígono). As fases de trabalho envolvidas no processo de vetorização dependem da estrutura de quem executa e dos meios disponíveis para tal.

Para fins de melhor clareza e entendimento será abordada aqui a metodologia adotada pela DSG. Segundo esta, a vetorização é composta por seis fases de trabalho pré-estabelecidas no Manual de Vetorização. As fases são: Digitalização matricial, Preparo, Confecção da moldura, Registro, Vetorização e Ligação. As seções a seguir abordarão cada uma dessas fases.

Digitalização matricial

Os originais de documentos cartográficos sejam impressos ou fotolitos, são imagens de tom contínuo. Isso significa que a transição entre tons e formatos é suave e uniforme. Infelizmente, os computadores não podem entender diretamente a

variabilidade harmônica de uma imagem de tom contínuo, pois processam apenas bits de informação. Mas o computador pode trabalhar com uma imagem se ela for transformada em pontos. A digitalização matricial é o processo pelo qual transformamos a imagem de tom contínuo em pontos por meio da conversão do meio analógico para o digital realizada em equipamento de varredura/captura de imagem denominado genericamente por *scanner*. Esses pontos criados durante a digitalização são mais adequadamente chamados de *pixels (picture elements)* e podem ser entendidos como sendo as células de uma matriz formada por linhas e colunas (surgindo daí o nome de digitalização matricial).

Na verdade os *pixels* são blocos de construção de imagens digitais, e devem ser considerados os seguintes fatores ao trabalhar com eles: tamanho, localização, valor e profundidade. A seguir será abordado cada um desses fatores.

a. Tamanho

O tamanho de um *pixel* determina a quantidade de informações numa digitalização. O tamanho de um *pixel é* determinado pela resolução geométrica na qual a imagem será digitalizada. Quanto menores os pixels, mais detalhes a imagem conterá.

b. Localização

Os *pixels* que compõem uma imagem são arrumados em uma matriz com eixos horizontal (linha) e vertical (coluna). Cada *pixel* tem uma localização específica e mensurável na matriz. Por exemplo, a localização de um *pixel* pode ser descrita como 250 *pixels* a partir da esquerda (coluna 250) e 300 *pixels* a partir de cima (linha 300). O número de células na matriz e, portanto, o número de *pixels* que compõem a imagem é determinado tanto pela resolução geométrica da imagem quanto por seu tamanho fisico. Quanto maior a resolução geométrica, maior o número de *pixels* em uma determinada área. Quanto maior o tamanho fisico da imagem, maior a área abrangida pela matriz.

c. Valor

Quando uma carta impressa é digitalizada, cada *pixel* recebe um valor baseado na cor do ponto original que representa. Todos os scanners coloridos adquirem informações de cor na forma de combinações variantes das cores aditivas, vermelho, verde e azul (RGB). Essas cores podem ser combinadas para representar virtualmente todo o espectro visível. Cada *pixel* pode representar

somente uma cor, mas quando muitos *pixels* pequenos de valores progressivamente diferentes são colocados perto uns dos outros, é criada a ilusão do tom contínuo.

d. Profundidade

A última característica de um *pixel é* a profundidade. Um *pixel* pode ter apenas um valor. A profundidade determina o número de possibilidades para o valor do *pixel.* Quanto maior a profundidade dos *pixels,* maior o número de dados necessários para armazenar a imagem. O mais importante para se ter em mente por enquanto é que, quanto maior for a profundidade do *pixel,* maior será o tamanho do arquivo.

As especificações técnicas usadas para a elaboração das normas de digitalização na DSG levam em conta três fatores básicos, a saber: os tipos de originais a serem digitalizados, os requisitos de qualidade para o tipo de trabalho a ser realizado e o consumo de tempo no processo produtivo. Esses fatores serão abordados nas seções seguintes.

Tipo de originais

Para fins de produção cartográfica é preciso que se tenha em mãos originais cartográficos em bom estado de conservação. Para este tipo de atividade existem basicamente dois tipos de originais em formato analógico, os quais são usados durante o processo de digitalização matricial, os diapositivos dos originais cartográficos em base plástica e a carta impressa em papel. Esses tipos de documentos possuem características distintas que impõem condições de processamento diferenciadas.

Requisitos de Qualidade

Para os dois tipos de originais apresentados é necessária que a imagem digitalizada corresponda às expectativas com relação à fidelidade de cor, uniformidade de tons e registro, bem como seja capturada em condições semelhantes às demais, na qual a capacidade de repetição do scanner, o foco e a resolução têm grande influência.

a) Fidelidade de cor

As cores que um *scanner* captura de um determinado documento cartográfico, como, por exemplo, uma carta impressa, depende da combinação de

sua fonte de luz, filtros, componentes eletrônicos e software. As imagens digitalizadas devem ter boa fidelidade nas diversas partes do espectro de cores para que possa representar adequadamente um documento em cores.

b) Uniformidade de tons

A transição entre tons está relacionada à faixa de tons que um *scanner* pode capturar. Para imagens com apenas dois tons (preto e branco), que é o caso dos diapositivos dos originais cartográficos, esta característica é irrelevante, entretanto para imagens obtidas em tons de cinza, que é o caso das cartas impressas, este fator se torna preponderante. A obtenção de uma imagem que reproduza fielmente as gradações de tonalidade é extremamente importante quando se pensa em trabalhar com imagem pancromática, pois nestas as cores são representadas por tonalidades de cinza distintas.

c) Registro

Um scanner captura as informações de imagens coloridas em canais separados compostos pela cor vermelha, verde e azul (RGB) e, em seguida, intercala-os no software. Se os três canais de cores não estiverem adequadamente alinhados, a imagem ficará fora de registro. Um registro pobre resulta em imagens borradas, com bandas de cor visíveis em torno dos cantos. O registro não precisa ser pobre para ser observável. Problemas de registro normalmente são resultantes da imprecisão do mecanismo que move o original ou o CCD.

d) Capacidade de repetição

A fonte de luz do scanner precisa de um tempo para se estabilizar. Quando a luz está ligada, ela pode variar em intensidade por um período de tempo que vai de alguns segundos a meia hora, até que atinja seu ponto de estabilidade. Em alguns modelos de scanner é necessário um tempo de espera para o seu uso. Isto permite que a fonte de luz se estabilize antes de ser iniciada a digitalização. Outros scanners ligam suas fontes de luz apenas quando a digitalização é feita. Esses scanners correm o risco de produzir as primeiras digitalizações diferentes das que são feitas depois, gerando resultados diferentes para o mesmo original.

e) Foco

O foco está relacionado à distância existente mantida entre o original e o CCD. Para que o foco não seja um problema na digitalização de originais é necessário que a superficie na qual o original é colocado esteja sempre à mesma distância do CCD que registra a luz. O foco é afetado por muitos fatores, tais como

o tipo de montagem usado, o grau de deformação do original e a quantidade de calor gerada pelo scanner.

f) Resoluções

O assunto resoluções pode parecer confuso, mas na verdade, é bastante objetivo. Ao contrário dos fatores cor e uniformidade de tons, que são subjetivos, a resolução é quantificável e, portanto, fácil de calcular.

Para a conversão de qualquer documento para o meio digital via digitalização matricial, bem como para originais cartográficos, é necessária a definição dos tipos de resoluções desejadas. Estas são a resolução geométrica (ou espacial), a resolução espectral e a resolução radiométrica.

As resoluções adequadas são questões fundamentais na digitalização, ainda que também sejam as menos compreendidas. A menos que a imagem seja digitalizada nas resoluções apropriadas, os resultados serão decepcionantes. Digitalizar documentos sem critérios de resoluções podem resultar em uma imagem grosseira e confusa. Por outro lado ao se superestimar a resolução pode se obter arquivos maiores que o necessário. Isso desperdiça tempo de digitalização e dinheiro, pois o custo para armazenar arquivos de imagens grandes é maior.

Resolução geométrica

A resolução geométrica ou espacial é simplesmente a definição do número de células *(pixel)* que o scanner deve criar para cada polegada linear do original. Por exemplo, ao digitalizar um original com 300 *dpi (dots per inch ou pixel* por polegada) produz-se 90 mil *pixel* (300 *pixel* horizontalmente x 300 *pixel* verticalmente) para cada polegada quadrada do original. Portanto, quanto maior a quantidade de *pixel* por polegada quadrada maior será a dimensão do arquivo de imagem.

Resolução espectral

A resolução espectral diz respeito à segmentação do espectro eletromagnético capturado pelo CCD durante a digitalização, ou em outras palavras, as faixas do espectro capturadas que serão usadas nos canais de cores do monitor do computador. Esta resolução para os dispositivos de varredura tipo scanner se restringem basicamente ao espectro visível, que pode ser segmentado em bandas individuais para formar imagens coloridas quando associados aos canais RGB do monitor ou segmentando todo o espectro visível numa única banda para formar uma imagem pancromática. A digitalização dos originais em bandas para

compor uma imagem colorida resultará em arquivos maiores, pois cada banda possuirá uma imagem em tons de cinza correspondente a faixa do espectro capturada, que juntas (três bandas) formará a imagem colorida.

Resolução radiométrica

Este tipo de resolução está diretamente ligado à quantidade de variação de tons de cinza ou níveis de cinza que um *pixel* pode assumir. Na prática, a adoção da resolução radiométrica se limita a alguns valores básicos tais como 1, 4 e 8 bits.

As especificações de resoluções adotadas no Manual de Vetorização da DSG para digitalização matricial dos diapositivos de originais cartográficos (fotolitos) e das cartas impressas, bem como os formatos de arquivos de saída, podem ser vistos na Tabela 2.4.

Base de dados analógica	Formatos de arquivo	Resolução Geométrica (min.)	Resolução Espectral	Resolução Radiométrica
Diapositivos e originais cartográficos	.tif, .rle	600dpi	-visível (P&B - preto e branco).	1bit
Folha impressa em papel	.tif, .rgb	300dpi	- visível (pancromático - tons de cinza); - visível (colorida - RGB).	1, 8 ou 24bits

Tabe 2.4 Especificaçoes técnicas para a digitalização matricial de originais cartograficos na DSG.
Fonte: Manual de Vetorização, DSG.

As resoluções apresentadas dizem respeito basicamente a aspectos de ordem prática. No caso dos originais cartográficos a resolução geométrica adotada para a digitalização é maior porque os arquivos gerados devem possuir uma quantidade maior de *pixels* para uma melhor definição dos elementos capturados. Isto se deve porque a técnica de vetorização em tela adotada para a conversão deste tipo de material é a semi-automática, motivo pelo qual também é adotado uma resolução radiométrica de 1bit. Já no caso de folhas impressas a resolução geométrica apresentada é menor, pois são previstos para este tipo de material, além da resolução de 1bit, outros tipos de resolução radiométrica as quais permitem somente a técnica de vetorização em tela manual.

Outro fator importante a ser comentado é que o scanner usado neste processo deve ser o de grande formato (padrão ISO AO), que permite gerar durante a varredura uma imagem única para cada carta evitando colagens ou mosaicos de imagens, com vistas a não ocorrer em possíveis erros oriundos deste tipo de procedimento.

Preparo

Esta fase tem por finalidade principal o preparo um exemplar impresso da carta a ser trabalhada, sobre a qual são indicados por meio de anotações quais os tipos de fonte e dimensões a serem usadas no processo de vetorização, bem como a indicação de detalhes para esclarecer quaisquer dúvidas em relação às feições das cartas que serão adquiridas. Tal fase visa a aplicação do padrão imposto pela Tabela de Aquisição de Dados, bem como, agilizar a execução do trabalho e eliminar possíveis enganos cometidos pelo operador encarregado do trabalho de vetorização.

Confecção de moldura

Esta fase, bem característica de uma instituição responsável pelo mapeamento sistemático, é responsável pela criação em ambiente de CAD do limite da área de desenho da carta a ser vetorizada. Por ser uma atividade específica que diz respeito a este tipo de mapeamento não serão comentados aqui os pormenores da operação. É interessante, porém ressaltar que a moldura é criada em formato vetorial com auxílio de aplicativos específicos, de acordo com a escala do mapeamento sistemático ao qual a carta pertence, e tem como objetivos limitar a área de desenho da carta, bem como, auxiliar na próxima fase do processo, que é o registro da imagem matricial.

Registro

Nesta fase é realizado o registro em relação a um referencial cartográfico dos arquivos matriciais com as informações da carta, por meio do procedimento conhecido como ajustamento. Para tanto, são realizadas observações sobre pontos homólogos de coordenadas conhecidas no sistema tomado como referência e no sistema a ser transformado. Estes dados são usados pa a calcular os parâmetros de transformação entre os dois espaços por meio de ajustamento pelo Método dos Mínimos Quadrados (MMQ) . O tipo de ajustamento aplicado é o paramétrico, que permite a partir de observações em dois sistemas distintos determinar, por meio da adoção de um modelo matemático, os valores de coeficientes de transformação [Gemael, 1994]. Estes coeficientes por sua vez estão associados a modelos matemáticos conhecidos que possibilitam o registro da estrutura matricial e posteriormente durante a vetorização das coordenadas dos pontos

que levarão as feições para o formato vetorial.

Para que se tenha uma vetorização dentro da precisão exigida, torna-se imprescindível que esta fase seja feita com o máximo rigor, pois esta parte do processo é uma das principais fontes de erro relativas ao posicionamento das feições espaciais da carta.

As precisões adotadas para o georreferenciamento levam em conta a aplicação do conceito de erro gráfico permitido, o qual estabelece que o erro natural cometido na observação de uma coordenada numa carta é da ordem de 0,2 mm na escala da mesma. O manual de Vetorização da DSG leva em consideração estas precisões e estabelece que o erro máximo permitido deve ser de metade do valor do erro gráfico. Para o caso das escalas do mapeamento sistemático pode-se então encontrar as seguintes precisões (Tabela 2.5):

Escala	Erro gráfico (m)	Erro Máximo Permitido (m)
1:250.000	50,00	25,00
1:100.000	20,00	10,00
1:50.000	10,00	5,00
1:25.000	5,00	2,50

Tabela 2.5 - Precisões estabelecidas para o georreferenciamento de cartas nas escalas do mapeamento sistemático.
Fonte: Manual de Vetorização da DSG

A cada georreferenciamento realizado são verificados os erros cometidos nas observações por meio da análise dos resíduos e do erro médio quadrático da média, para constatar se o grau de precisão estipulado no Manual de Vetorização está sendo atendido.

No contexto das instituições de mapeamento são registrados basicamente dois tipos de documentos em formato matricial, já mencionados anteriormente: os arquivos matriciais dos diapositivos de originais cartográficos e das cartas impressas. A seguir serão detalhados os procedimentos executados para cada um dos tipos mencionados.

Originais Cartográficos

Cada original após a conversão analógica digital e antes de ser submetido à vetorização, é submetido ao procedimento de registro. Para tanto, são coletadas as coordenadas dos cantos de ajuste dos originais em formato matricial e

as coordenadas de canto da moldura da folha gerada por aplicativo específico em formato vetorial. O modelo matemático adotado para realizar a transformação das coordenadas de *pixel* para o sistema de coordenadas de projeção cartográfica é o Afim (de 1ª ordem). A escolha de tal modelo é baseada no fato do mesmo ser matematicamente linear e por possibilitar a modelagem das distorções dentro da precisão estipulada. A formula matemática que representa este modelo matemático [Wolf, 1983] é dado em formato matricial por:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}$$

onde,

$\begin{bmatrix} X \\ Y \end{bmatrix}$ - é uma matriz coluna (vetor) contendo as coordenadas *X* e *Y* no sistema de projeção cartográfica;

$\begin{bmatrix} x \\ y \end{bmatrix}$ - é uma matriz coluna (vetor) contendo as coordenadas *x* e *y* no sistema de coordenadas de pixel;

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ - é uma matriz quadrada (2x2) contendo os parâmetros de transformação *a, b, c* e *d*, relativos a rotação, a fatores de escala e a condição de não ortogonalidade entre eixos;

$\begin{bmatrix} e \\ f \end{bmatrix}$ - é uma matriz coluna (2x1) contendo os parâmetros de transformação *e* e *f* relativos às translações nos eixos x e y respectivamente.

Os significados dos parâmetros *a, b, c e d* são dados pelas seguintes expressões matemáticas:

$$a = S_x \frac{\cos\alpha}{\cos\alpha} \qquad b = S_y \frac{\operatorname{sen}\alpha}{\cos\alpha}$$

$$c = S_x \frac{\operatorname{sen}\alpha}{\cos\alpha} \qquad d = S_y \frac{\cos\alpha}{\cos\alpha}$$

onde,

S_x e S_y - são os fatores de escala em x e y respectivamente;

? - é o ângulo de rotação; e

? - é a condição de não ortogonalidade entre eixos.

A manipulação das expressões apresentadas acima permite que se chegue aos valores dos fatores de escala em x e y (S_x e S_y), o valor da rotação (?) e o valor da condição de não ortogonalidade entre eixos (?), que matematicamente podem ser expressos por:

? ? $\tan^{?1}\left(\frac{? ? c ?}{? d ?}\right)$

? ? ? ? $\tan^{?1}\left(\frac{? b ?}{? a ?}\right)$? ? ? ? ? ? ? ? ? ?

S_x ? a ? $\frac{\cos ?}{\cos(? ? ? ?)}$

S_y ? d ? $\frac{\cos ?}{\cos ?}$

O modelo matemático apresentado exige que sejam realizadas observações sobre pelo menos três pontos com coordenadas conhecidas nos dois sistemas considerados, o de CAD e o da imagem em formato matricial. Isto permite uma solução única na determinação dos parâmetros de transformação. A leitura de quatro pontos homólogos (cantos da moldura e cantos de ajuste dos originais) garante a condição mínima para determinação dos parâmetros de transformação, fornecendo somente dois graus de liberdade ao sistema, os quais possibilitam um controle de qualidade mínimo para o procedimento de registro.

A experiência adquirida em trabalhos de vetorização na DSG, entretanto demonstrou que a metodologia estabelecida no manual de vetorização que prevê a adoção dos cantos da moldura como pontos observáveis no ajustamento para a determinação dos parâmetros de transformação, pode ocasionar sérios erros de posicionamento das feições no interior da carta. Isto se deve porque alguns originais possuem distorções de difícil detecção e às vezes de natureza desconhecida. Embora esse problema possa ser detectado em alguns originais, o uso deste tipo de material para conversão de base analógica para o meio digital, não deve ser descartado, pois o

mesmo apresenta vantagens não encontradas nos arquivos matriciais de cartas impressas. Os vetores oriundos do processo de digitalização vetorial do arquivo matricial dos originais cartográficos apresentam traçado mais bem delineado e preciso, que aliado ao fato de permitirem uso de técnicas de vetorização automática e semi-automática, que diminuem consideravelmente o tempo de conversão, tomam este tipo de material ideal para trabalhos desta natureza. Os principais casos, verificados na prática, que acarretam erros de georreferenciamento de originais cartográficos usando as marcas de registro são:

- As quatro marcas de canto encontradas nos originais embora coincidentes não garantem a consistência das representações cartográficas no interior da carta. Isto pôde ser constatado em algumas ocasiões onde foram obtidas perfeita sobreposição das marcas de registro dos originais e as feições internas das cartas apresentaram deslocamentos inaceitáveis;
- A base plástica dos originais não pode ser considerada estável e isenta de distorções. Neste caso as inconsistências no desenho cartográfico tendem a se apresentar em regiões distintas. Sua constatação é possível por meio de inspeção visual, sobrepondo-se os originais em base analógica sobre uma mesa de luz com tampo de vidro; e,
- O scanner usado na digitalização matricial dos originais pode apresentar problemas devido à falta de manutenção ou calibração. Neste caso a constatação das inconsistências será visível somente em meio digital ao abrir os arquivos dos originais em formato matricial em software apropriado.

Quando ocorre o primeiro caso é usado um procedimento alternativo para o georreferenciamento. Como os originais apresentam somente os quatro cantos de registro como informação para uso no georreferenciamento (exceção à regra o original de cor preta) é necessária a adoção de uma base de referencia intermediária que seja geometricamente confiável e possibilite o registro dos originais. A idéia consiste em criar uma base sobre a qual seja possível coletar pontos diferentes daqueles mencionados anteriormente. Para tanto a carta impressa (produto da soma dos originais de cores sépia, azul, verde, vermelha e preta) é submetida à digitalização matricial e georreferenciamento, tomando-se por base o quadriculado

em coordenadas UTM da folha gerado por aplicativo específico em formato vetorial. O uso do quadriculado permite que seja coletado um número maior de pontos de controle possibilitando que haja uma maior controle das distorções existentes na carta impressa. A carta impressa usada como base, por sua vez, além de possibilitar a escolha de uma maior quantidade de pontos de controle para realizar o registro dos originais com distorções, permite que os pontos de canto de registro sejam descartados do processo. Podem ser escolhidos quaisquer detalhes das feições como pontos de controle para serem usados no georreferenciamento dos originais. Alguns exemplos seriam: interseções de estradas (originais de cor preta), bifurcações de rios (originais de cor azul), e assim por diante.

No segundo caso a solução se assemelha àquela usada para a primeira situação. Da mesma maneira o artificio de se gerar uma referência baseada no arquivo matricial da carta impressa é adotado. O diferencial se dá, porém no tratamento das distorções, pois devido à regionalização destas no espaço geográfico da carta se faz necessário um tratamento diferenciado para as mesmas. Neste caso é realizada uma extração da região geográfica que apresenta distorções gravando-se esta num novo arquivo matricial, o qual é registrado separadamente. Dependendo do grau de distorção encontrado o modelo matemático a ser adotado pode ser mudado, ou seja, pode se optar por modelos matemáticos não lineares. Contudo cabe frisar que esta alternativa é viável somente em regiões não muito extensas e nas quais se tenha uma abundância de pontos que permita uma boa modelagem das distorções. Em situações normais o uso desta técnica não é aconselhável, pois além de exigir grande quantidade de pontos pode ocasionar erros sérios de posicionamento em locais onde não existam pontos de referência.

O terceiro e último caso a solução mais adequada seria realizar uma nova digitalização dos originais após manutenção do equipamento de digitalização matricial.

Caso não seja possível este procedimento pode-se aplicar alternativas não usuais para a correção dos originais. As soluções a serem adotadas neste caso dependem basicamente da análise do tipo de problema e devem ser precedidas ao registro. Por exemplo, se o arquivo matricial apresentar deslocamentos uniformes no sentido das colunas ou das linhas da matriz formada pela varredura do dispositivo digitalizador é possível corrigi-los num software de tratamento de imagem, modificando a estrutura da matriz de forma a coincidir as linhas ou colunas

deslocadas. Outro exemplo seria a remoção de ruídos criados durante a digitalização matricial por meio do uso de filtros especiais. Deve ser ressaltado, porém que a adoção das operações mencionadas para as situações apresentadas logo acima podem não resultar na correção esperada, neste caso deve-se abandonar este tipo de tratamento, e a única alternativa seria a apresentada anteriormente. O registro dos originais corrigidos por meio desta alternativa deve ser executado com uma atenção redobrada, pois se deve ter em mente que a estrutura matricial original construída pelo *scanner* foi corrigida por meio de intervenção humana.

Do que foi apresentado, para os dois primeiros casos, chega-se à conclusão que urna verificação prévia dos originais é imprescindível, pois por meio dela é possível se constatar:

- a integridade física dos originais;
- o estado de limpeza da superficie dos originais;
- o estado das informações dos originais;
- a coincidência ou não das marcas de registro dos originais;
- a existência de regiões onde o original possa ter sofrido distorções;

Conclui-se também, para a última situação, que o estado de conservação do dispositivo de conversão analógico-digital, é um fator preponderante, pois dela depende a qualidade das imagens geradas. Portanto a manutenção (limpeza e calibração) do equipamento deve ser realizada periodicamente e de acordo com as recomendações do fabricante.

Carta Impressa

Assim como o original, a carta impressa após conversão para o meio digital, via digitalização matricial, também é submetida ao georreferenciamento. Ao contrário da situação restritiva existente nos originais que a princípio impossibilitam a coleta de mais de quatro pontos de controle (cantos de registro do original), o georreferenciamento do arquivo matricial da carta impressa permite a coleta de mais pontos, pois como comentado anteriormente é possível recriar o quadriculado UTM da carta por meio de software e usá-lo para ajustar as coordenadas de *pixel* existente no arquivo matricial. Para tanto são selecionados sobre a imagem da carta as coordenadas dos pontos das interseções do quadriculado UTM existente e seus pontos homólogos sobre os vetores. Entretanto,

para manter uma padronização de procedimento é adotado como pontos observáveis, num primeiro momento, somente os quatro cantos da moldura. Caso as precisões para o georreferenciamento não sejam atingidas aí se procede à leitura de outros pontos. O modelo matemático adotado, assim como no caso anterior é o Afim de 1 a ordem. A determinação dos parâmetros de orientação deste, quando utilizados mais pontos além daqueles que identificam os quatro cantos da carta, são melhores obtidos. Isso é possível porque a existência de pontos em abundância permite uma análise de resíduos mais abrangente a qual possibilita inferir parâmetros para a eliminação de pontos espúrios. O uso do quadriculado UTM além de possibilitar maior coleta também permite uma verificação visual da coincidência dos vetores e com a imagem da carta.

Vetorização

A eficiência desta fase depende da qualidade do software de vetorização e da experiência do operador. A atividade de digitalização vetorial é uma tarefa muito cansativa, normalmente consome muito tempo e em decorrência deste fator alguns erros podem ser cometidos. Por isso, os softwares usados para este fim geralmente fornecem ferramentas que auxiliam o operador a identificar e corrigir possíveis erros grosseiros.

Basicamente existem duas técnicas para a conversão em formato vetorial de dados espaciais encontrados em cartas impressas ou originais cartográficos: a baseada na vetorização em tela e a vetorização manual. A diferença básica entre elas é que enquanto a primeira depende exclusivamente do uso de softwares a segunda depende, além destes do dispositivo conhecido como mesa digitalizadora.

Vetorização em tela

A vetorização em tela pode ser realizada sobre originais cartográficos ou cartas impressas, transformadas em imagens por meio de conversão analógico-digital via *scanner*. A primeira situação é a adotada pelas instituições de mapeamento nacional para passar para o meio vetorial cartas já existentes visando a atualização ou simplesmente o uso de algumas informações existentes tais como curvas de nível e pontos cotados. Já a segunda opção é adotada pela maioria dos usuários,

justamente por não possuir como fonte de dados os originais.

Vetorização de originais cartográficos

Na digitalização vetorial dos arquivos dos originais convertidos, são empregadas técnicas de vetorização semi-automática as quais permitem obter um melhor traçado das feições e a redução do tempo de produção. Essas técnicas estão associadas à aplicação de algoritmos de varredura que usam arquivos matriciais com resolução de 1bit (preto e branco) para localizar as feições a serem vetorizadas. Esta resolução permite ao algoritmo interpretar o espaço bidimensional do arquivo matricial como um espaço composto por células (pixel) que representam a informação de forma lógica pela representação do número 1 ou 0, em outras palavras, com informação ou sem informação respectivamente. O espaço criado por essas condições é conhecido como "booleano", e sua existência é fundamental para SIG que trabalham com estrutura matricial. Esta característica peculiar facilita a aplicação de rotinas que tem como prerrogativas percorrer somente células que contenham informação (valores lógicos igual a 1) que implementados para se posicionar de acordo com a célula central da linha construída permite a automatização do processo de vetorização. Nestas condições é possível se obter grande precisão no traçado das feições, bem como, agilizar o processo produtivo por meio da redução do tempo de trabalho.

A técnica envolvida prevê ainda que a digitalização vetorial dos originais seja realizada individualmente, de acordo com os temas apresentados pelas cores usadas para gerar a impressão *offset*, o que de certa forma já pode ser considerada uma pré-estruturação das informações em temas.

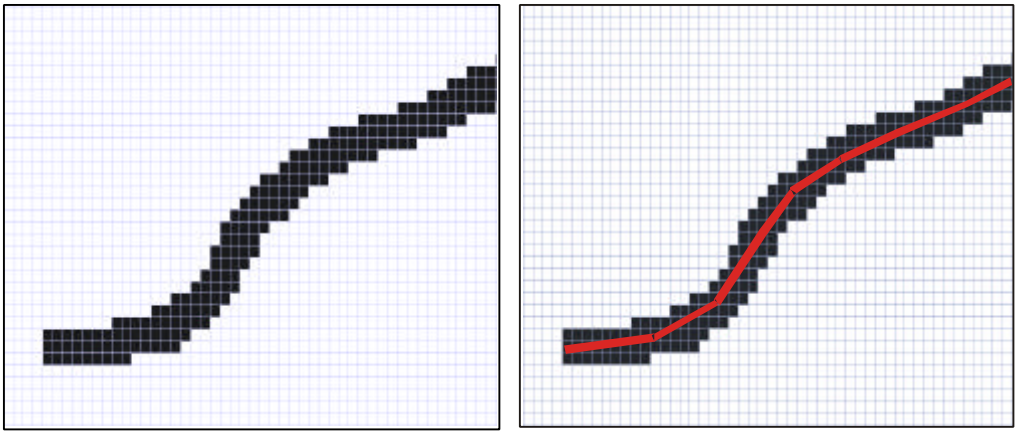


Figura 2.22 - Detalhe da estrutura matricial com resolução de 1 bit, princípio da formação do espaço "booleano".

As cores usadas para os originais cartográficos são: verde, sépia, azul, preto e vermelho. A cor verde representa na carta as feições de vegetação, tais como florestas, culturas, etc. A cor sépia por sua vez representa feições relacionadas basicamente a altimetria e sua toponímia, tais como curvas de nível, pontos cotados, bem como terrenos rochosos ou arenosos, representando ainda feições complementares da planimetria como barragens de açudes. A cor azul tem como finalidade representar as feições ligadas a hidrografia e sua toponímia, tais como rios, lagos, linhas de costa, etc. A cor vermelha tem como principal finalidade salientar as feições criadas artificialmente e de importância estratégica para as atividades humanas tais como sistema viário permanente e regiões densamente povoadas como áreas urbanizadas das grandes cidades. Por fim o preto representa, de forma geral, as feições de planimetria, tais como, construções significativas, contorno das estradas, arruamentos, pistas de pouso, etc, além de informações marginais e de moldura, tais como legendas, coordenadas geográficas e de projeção cartográfica UTM.

Após o georreferenciamento dos originais o trabalho de conversão se resume basicamente em capturar as feições por meio do emprego das ferramentas do software de vetorização de acordo com a Tabela de Aquisição de Dados usada no projeto.

Vetorização de cartas impressas

Quando não se possui o original cartográfico como fonte de dados a solução para conversão analógica digital se dá por meio do uso das cartas impressas em substituição aos anteriores. Neste caso a técnica de vetorização a ser usada deve ser a manual, pois os arquivos das cartas impressas impossibilitam o uso de técnicas semi-automáticas.

Como visto na fase de Digitalização Matricial são usadas para as cartas impressas resoluções radiométricas de 1, 8 e 24 bits. Essas resoluções são aplicadas de acordo com a existência de fatores condicionantes tais como: tipo de scanner usado, software disponível e espaço em disco.

O espaço em disco rígido é um ponto importante que deve sempre ser considerado. Por exemplo, um arquivo de imagem de uma carta em 1 bit é extremamente pequeno quando comparado a um de 8bits, que por sua vez é menor

que um de 24 bits. Apesar da existência hoje de discos rígidos com grande capacidade de armazenamento é importante lembrar que imagens coloridas de 24 bits de cartas topográficas consomem um espaço considerável de memória.

A resolução de 1 bit, já comentada anteriormente, fornece uma imagem em preto e branco passível de aplicação de técnicas de vetorização semi automática. Porém para o caso de arquivos de cartas impressas, as vantagens desta técnica são anuladas. O fator que impõe esta restrição está relacionado à quantidade de cruzamentos de linhas de temas diferentes existente na carta impressa, os quais quando levados para a estrutura matricial confundem o algoritmo de busca causando grande quantidade de erros grosseiros no momento da vetorização.

A resolução de 8 bits corresponde a uma imagem da carta em tons de cinza, também denominados níveis de cinza. O valor máximo e mínimo destes níveis é representado respectivamente por 255 que corresponde ao branco e 0 que corresponde ao preto (variação total de 256 níveis). A vantagem ao se trabalhar com este tipo de resolução para os arquivos matriciais de cartas reside basicamente na representação da imagem de forma contínua, o que facilita a execução da vetorização por possibilitar maior definição das feições permitindo maior discernimento por parte do operador. Um arquivo matricial nestas condições, entretanto, não permite a aplicação de técnicas de vetorização semiautomática, pois como visto anteriormente os algoritmos de buscas são desenvolvidos para capturar informações sobre imagens em preto e branco (1 bit).

A resolução de 24 bits por sua vez é oriunda da transformação da imagem capturada por scanner em três imagens distintas em tons de cinza, cada qual com as informações referentes aos canais de cores RGB e com resolução de 8 bits, que associadas ao canhão, vermelho *(Red)*, verde *(Green)* e azul *(Blue)*, existentes num monitor formam uma imagem colorida. A vantagem deste tipo de arquivo matricial reside no fato do mesmo ser uma reprodução fiel da carta impressa com detalhes representados em cores conforme aparecem na base analógica. Este fator permite que **o** operador consiga identificar facilmente **os** detalhes representados na carta anulando praticamente qualquer problema de interpretação que porventura possa se apresentar.

O registro das informações contidas no arquivo da carta, como comentado anteriormente na fase de Georreferenciamento, pode ser mais bem

acurado que nos originais.

O trabalho de vetorização, assim como na técnica anterior mencionada, se resume basicamente em extrair as feições da carta por meio da digitalização vetorial das feições existentes. Entretanto, como não é possível aplicar a técnica de vetorização semiautomática o rendimento do trabalho e a qualidade do produto final, ficam prejudicadas.

Vetorização manual (com mesa digitalizadora)

A técnica de digitalização vetorial por meio de mesa digitalizadora é caracterizada pela extração das informações vetoriais diretamente das cartas analógicas. Foi a primeira técnica de conversão de cartas para o meio digital em formato vetorial. Hoje é muito dificil encontrar este tipo de dispositivo em uso, pois com o advento dos scanners de grande formato praticamente todos os produtores de cartas migraram para a digitalização vetorial em tela. Entretanto é interessante tecer alguns comentários em relação às mesas digitalizadoras por ter sido estas desenvolvidas basicamente para uso em atividades de reprodução cartográfica.

A aquisição de dados por meio de mesa digitalizadora é realizada por um dispositivo de ponteiro tipo mouse o qual é movimentado pelo operador de maneira a percorrer o traçado das feições diretamente sobre a carta impressa de modo o mais fiel possível. O que está por trás desse sistema é um dispositivo de codificação de movimento num plano (plano da mesa) baseado numa trama de sensores eletrônicos dispostos lado a lado no sentido do comprimento e da largura da mesa, de maneira a representar o espaço bidimensional em termos de um par de eixos cartesianos (x e y). Esses sensores formam basicamente uma trama ortogonal de fios embutidos que são sensibilizados ao se passar um cursor sobre eles. Este cursor é o mouse que possuí uma espécie de bobina capaz de gerar um campo magnético passível de ser percebido pela trama de sensores. O movimento deste sobre a mesa é entendido como variações em termos de pulsos elétricos que podem ser convertidos na medição da posição do cursor sobre a mesa [Dale & McLaughlin, 19901. Os valores registrados (coordenadas da mesa) dependerão das características da mesa digitalizadora.

Para transformar as coordenadas da mesa em coordenadas no sistema plano de referência do mapa, é necessária a digitalização de pontos como os cantos

do mapa ou interseções entre linhas, para que se proceda à determinação dos parâmetros de transformação matemática entre os sistemas. Esta fase de determinação destes parâmetros consiste no georreferenciamento das coordenadas dos pontos digitalizados. A precisão de posicionamento que se pode chegar com o uso de mesas digitalizadoras está diretamente relacionada à maneira como são construídas, ou seja, de acordo com o espaçamento entre os fios da trama. Pode-se obter precisões de construção da ordem de 0,01mm, porém a exatidão da digitalização não chega a este nível, principalmente devido a fatores humanos. A digitalização numa mesa digitalizadora é um processo extremamente tedioso e exige concentração por parte do operador. Duas horas de esforço concentrado é o máximo que a maioria das pessoas consegue trabalhar sem interrupção com este equipamento [Dale & MacLaughlin, 1990].

Aspectos importantes

Segundo o Manual de Vetorização, tanto a técnica baseada na vetorização em tela e a de vetorização manual devem seguir o mesmo fluxo de trabalho. Este prevê que sejam usados padrões pré-definidos para aquisição das feições encontradas na carta, por meio de arquivos que possuam:

- Tipos de fonte: este arquivo deve conter todas as fontes previstas na fase de Preparo, para que se possa reproduzir fielmente a toponímia encontrada na carta;
- Tipos de símbolos para feições pontuais: este arquivo deve conter os símbolos cartográficos encontrados na carta para representar de forma visual a informação visual capturada;
- Tipos de vetores para feições lineares e de área: este arquivo contém a codificação para as feições encontradas na carta, geralmente por meio de atributos lógicos encontrados em aplicativos tipo CAD, servindo para individualizar as feições.

Antes de realizar a vetorização propriamente dita de uma determinada feição linear ou de área, quando se usa um software de vetorização, deve-se configurar os parâmetros necessários para que a linha vetorial reproduza fielmente a linha matricial. Estes softwares geralmente fornecem três tipos de modalidade de digitalização vetorial, a saber:

- **Automática:** nesta modalidade a linha matricial é vetorizada automaticamente sem que seja necessário o operador definir a direção de vetorização. Basta que seja selecionado o ponto inicial do elemento a ser vetorizado e ela é vetorizada até o seu final ou até encontrar algum obstáculo (representado geralmente por uma interseção). Esta modalidade pode ser usada para feições lineares ininterruptas e sem interrupções, como por exemplo, as curvas de nível;
- **Semi-automática:** nesta modalidade, para linha matricial ser vetorizada é necessário que o operador defina a direção da vetorização. Basta selecionar dois pontos de uma linha matricial definindo uma direção para que a mesma seja vetorizada até o seu final ou encontrar algum obstáculo (interseção). Esta modalidade, que difere da anterior pela interferência do operador, pode ser usada para todas as feições lineares e de área;
- **Manual:** nesta modalidade, a linha matricial é vetorizada ponto a ponto por meio da construção de uma rede de pontos manualmente. Para aplicação desta modalidade é necessário somente percorrer a linha matricial com o cursor criando os pontos um a um. Esta modalidade difere das anteriores justamente por depender inteiramente da ação humana, e pode ser aplicada para vetorização de qualquer feição linear ou de área.

Outros fatores importantes para serem definidos quando se usa a modalidade de vetorização automática ou semi-automática em originais cartográficos são os seguintes:

a. Comportamento do software no encontro de obstáculos:

Os obstáculos encontrados no momento da vetorização são geralmente representados pelas interseções entre linhas matriciais. A ferramenta de digitalização nesses locais precisa de uma definição por parte do operador para saber qual o tipo de comportamento deve assumir. O comportamento mais adequado para estas ferramentas é que no momento do encontro do obstáculo o software pare e aguarde a definição da direção a ser tomada por meio da iteração com o operador. Isto é comum para a vetorização de linhas que compõe sistemas de rede tais como

hidrografia e malha viária. Na vetorização de elementos lineares contínuos como curvas de nível pode-se definir para a ferramenta que o comportamento seja o de percorrer as linhas sem parar nas interseções seguindo sempre a mesma direção evitando a mudança de direção de forma abrupta.

b. Configuração para vetorizar linhas descontínuas:

Algumas feições dos originais cartográficos são formadas por linhas matriciais de traçado descontínuo, por exemplo, curvas de nível auxiliares, rios intermitentes, dentre outras. Para a vetorização destes elementos se faz necessária à definição do espaçamento de *gap*, ou seja, o salto entre dois pontos que a ferramenta deve assumir no momento da vetorização destes tipos de linhas matriciais. Os valores dos espaçamentos recomendados no manual de Vetorização da DSG são apresentados na Tabela 2.6.

Escala da Carta	Espaçamento (m)
1:25.000	50
1:50.000	60
1:100.000	100
1:250.000	200

Tabela 2.6 - Valores de *gap* para vetorização de linhas matriciais descontínuas.
Fonte: Manual de Vetorização da DSG

c) Suavização e Filtragem de vértices:

Estes parâmetros são usados para suavizar e reduzir vértices no momento da vetorização. A suavização é utilizada para tornar suave, dentro de uma dada tolerância, os vértices que tenham um ângulo muito acentuado. A filtragem, que também se baseia numa tolerância adotada previamente, é utilizada para a eliminação de vértices criados excessivamente pela ferramenta de vetorização. Este fator é importante, pois dele dependerá o tamanho final do arquivo vetorizado e a fiel representação de feições vetorizadas.

Procedimentos adotados na vetorização

a. Procedimentos gerais

Tendo em vista a posterior validação geométrica do arquivo vetorial, recomenda-se que, durante a vetorização, crie-se "nós" nas interseções das linhas. Cabe ressaltar que a adoção deste procedimento, "nós" também devem ser criados nas interseções das linhas do desenho com as linhas que definem a moldura da carta, fazendo com que esta fique secionada em diversas partes. Este fator é extremamente importante para a fase de validação, pois por meio deste procedimento à validação dos elementos de área será facilitada.

Existem nos originais cartográficos feições representadas por tramas ou retícula, por exemplo, área sujeita a inundação, terreno rochoso, dentre outros. Nestas regiões é necessária a aplicação da modalidade de vetorização manual, pois as referidas tramas ou retículas não permitem ou dificultam a vetorização automática ou semi-automática. Para a vetorização de alguns elementos com trama, entretanto, tais como, rios de margem dupla e lagos permanentes, em princípio, é possível a vetorização semi-automática, tendo em vista que a retícula de massa d'água presente no original, normalmente, não é capturada pelo scanner na digitalização matricial com resolução geométrica de 800 dpi.

A reprodução de símbolos cartográficos existentes no arquivo matricial da carta deve ser realizada por meio do posicionamento correto da célula prevista. Estas células são construídas em ambiente de CAD numa determinada escala, sendo, portanto necessário aplicar um fator de escala para a perfeita reprodução do símbolo representado na carta. As células usadas no processo de confecção cartográfica da DSG foram criadas com uma escala compatível a 1:25.000, sendo portanto necessário a aplicação de fatores multiplicadores de escala para representá-las em outras escalas do mapeamento sistemático. A Tabela 2.7, a relaciona os fatores de escala usados pela DSG.

A inserção de toponímias encontradas nas cartas, em geral, deve ser feita de maneira que o texto não apresente rotação, mesmo que o original esteja inclinado.

Escala da Carta	Fator de Escala
1:25.000	1
1:50.000	2
1:100.000	4
1:250.000	10

Tabela 2.7 - Valores de ampliaçao de escala
para as células usadas pela DSG.
Fonte: Manual de Vetorização da DSG

b. Vetorização da Hidrografia

Consiste na vetorização dos elementos (lineares, pontuais e de área), bem como na colocação dos textos e símbolos encontrados na carta referentes ao original de hidrografia (cor azul). Basicamente o trabalho de vetorização é dividido em quatro partes:

- Lagos, rios e canais permanentes;
- Rios e canais intermitentes;
- Elementos de área que possuam trama; e,
- Colocação de células e textos da hidrografia.

Para garantir a existência de um "nó" em cada ponto de confluência dos rios, e maior espaço percorrido pela ferramenta de vetorização, deve-se vetorizá-los no sentido de jusante para montante, e com a opção de parar quando encontrar uma interseção.

Para a vetorização de elementos tais como recifes, cursos d'água canalizada, e outros similares, é recomendado o uso da modalidade de vetorização manual, pois os referidos elementos lineares possuem tramas que dificultam a aplicação da modalidade semi-automática.

A inserção das células representando corredeiras e rápidos não representáveis em escala deve ser executada utilizando-se as células respectivas a essas representações.

As áreas de inunda, brejo, arrozal, devem ser vetorizadas de acordo com o seu contorno, por meio de vetorização manual uma vez que tais áreas não possuem linhas de contorno aparente.

Devem ser vetorizadas as duas margens dos rios de margem dupla.

Os açudes devem ficar abertos nos locais onde estão definidas as posições das barragens.

c. Vetorização da Altimetria

Consiste na vetorização dos elementos (lineares, pontuais e de área), bem como na colocação dos textos e símbolos encontrados na carta referentes ao original de altimetria (cor sépia). Basicamente o trabalho de vetorização é dividido em três partes:

- Curvas de níveis mestras e simples;
- Curvas aproximadas e curvas de forma;
- elementos de área que possuam trama ou retícula; e,
- Colocação das células e textos da altimetria.

A vetorização de linhas extensas e densificadas como, por exemplo, curvas de níveis sinuosas de regiões serranas devem ser executadas com o máximo cuidado, pois a vetorização semi-automática pode se enganar e saltar de uma curva para outra.

A inserção de pontos cotados (células) deve ser feita de maneira rigorosa para que a o ponto seja posicionado corretamente. Deve-se adotar a célula correta para os diversos pontos cotados encontrados no original de altimetria.

Para a vetorização de curvas de nível mestra deve-se ter cuidado nos locais onde as mesmas são interrompidas pelo texto que da o valor de altitude da curva, pois caso contrário é possível que o mesmo seja vetorizado como parte integrante da do elemento linear.

Para a vetorização de elementos tais como afloramento rochoso, praia, dunas, e outros similares, é usada a modalidade de vetorização manual, pois os referidos elementos de área não possuem linhas delimitadoras impossibilitando a aplicação da modalidade semi-automática.

As cotas das curvas mestras devem ser colocadas sobre as curvas, sem interrompê-las.

Nas colocações dos textos respectivos a cota deve-se colocar a origem do texto na mesma origem das células, a fim de facilitar a leitura dos valores das mesmas em rotinas criadas para futura conversão de 2D para 3D.

d. Vetorização da Planimetria

Consiste na vetorização dos elementos (lineares, pontuais e de área), bem como na colocação dos textos e símbolos encontrados na carta referentes ao

original de planimetria (cor preta). Basicamente o trabalho de vetorização é dividido em três partes:

- Vetorização do traçado viário;

- Colocação de células de edificações e infra-estrutura; - Vetorização de área urbanizada;

A vetorização da planimetria deve ser executada de maneira criteriosa, pois os traçados viários encontrados nas cartas e originais são extremamente importantes em temos posicionais. Portanto, o perfeito casamento do posicionamento dos vetores com as feições encontradas no arquivo matrical devem ser realizados com rigor.

Parte III – QUALIDADE DE DADOS EM ESTRUTURA VETORIAL PARA USO EM SIG

2.3 Qualidade de dados para SIG

Os SIG são ferramentas poderosas que exigem investimentos relativamente alto quando comparado a ferramentas de desenho como o CAD. O custo alto advém não somente da aquisição de hardware, software e treinamento de pessoal, mas também dos não considerados custos de aquisição e tratamento de dados. Os custos atrelados a estas atividades posteriores podem aumentar consideravelmente o valor final de um projeto. Em alguns casos o volume de dados a ser tratado é considerável e pode levar a custos operacionais elevados que geralmente são embutidos no custo final de um projeto podendo desencorajar a execução do mesmo.

Qualquer atividade envolvendo o emprego de um de SIG tem embutido por trás o controle de qualidade dos dados. Este controle é a única maneira de se garantir uma maior probabilidade de acerto no processo de tomada de decisão e tem por finalidade reduzir ou até mesmo eliminar os possíveis erros de entrada de dados.

O processamento de dados geográficos por sua vez pressupõe que os dados usados na modelagem cartográfica estão totalmente livres de erros. Entenda-se “livre de erros” como não somente livre de erros causados por faltas cometidas no levantamento, mas também erros estatísticos, que provocam a livre variação dos mesmos. Em outras palavras, ao se realizar, por exemplo, uma sobreposição de mapas espera-se que ambos os mapas possam ser tratados como documentos possuidores de padrão uniforme de qualidade de dados.

A qualidade de dados é um fator que preocupa a maioria dos usuários de SIG por ser uma constante nas atividades desenvolvidas com estes sistemas. Profissionais ligados á área de produção cartográfica tem o conceito de qualidade de dados bem definido, pois sabem que o traçado de algumas feições sobre o mapa sofre algumas generalizações

para serem representadas. O contorno de áreas de florestas e o traçado da linha de praia encontrado nos mapas podem servir como exemplos da qualidade de dado quando comparado com o traçado de uma estrada.

O conhecimento da qualidade dos dados é crítico para realizar o julgamento dos tipos de aplicações para os quais eles são apropriados. Em análises de dados em meio analógico erro no desenho das feições dos mapas podem ser facilmente corrigidos ou até mesmo desconsiderados, mas em ambiente digital, que trata os dados logicamente, estes mesmos erros podem causar uma grande diferença no resultado ou até mesmo impedir a execução do processo de análise. A verificação e correção destes dados em ambiente digital requerem processamento extra que podem ser realizados por automação de processos e ou manualmente por iteração com o operador do sistema.

O custo de verificação dos dados é diretamente proporcional ao grau de rigor estabelecido, ou seja, quanto mais rigoroso o procedimento de verificação de qualidade maior é o custo. Este custo por sua vez abrange não somente o resultado dos gastos da execução do procedimento, mas também e principalmente, os atrasos causados no processo de produção devido à execução do procedimento e correção dos erros. A padronização dos procedimentos necessários para verificar a qualidade bem como os processos a serem usados nas correções são uma maneira de manter o mesmo grau de correção para os mapas digitais.

Infelizmente, para a maioria dos usuários, a qualidade dos dados geográficos é geralmente examinada somente após terem sido tomadas decisões incorretas e as perdas financeiras já ocorreram.

O aumento de produtores de informação geográfica e do número de usuários está fazendo com que diversos produtos de natureza espacial sejam encontrados com deficiências, tais como, erros e pobreza de desenho. Outro fator importante é que os dados gerados por seus idealizadores, às vezes com propósitos bem definidos, são usados indiscriminadamente para qualquer aplicação.

A obtenção indiscriminada de dados espaciais para SIG baseado na conversão de material cartográfico existente é um bom exemplo para ilustrar este tipo de prática. Caso esta conversão seja realizada indistintamente sem uma análise prévia do material existente e

um planejamento de tratamento adequado dos vetores adquiridos podem surgir problemas que provavelmente irão prejudicar a usabilidade de tais sistemas.

2.3.1 Entraves ao uso do SIG

Considerando que o melhor emprego de um SIG esta relacionada diretamente a tomada de decisão sobre dados espaciais, pode-se dizer que existem dois entraves para a otimização de seu uso:

a. Cultura cartográfica

A maioria dos pacotes de SIG utilizados pressupõe que o usuário tenha bons conhecimentos de cartografia e esteja acostumado a trabalhar com mapas, porém o que ocorre na prática é que a maioria dos usuários de SIG ainda se encontra num estágio pouco avançado em relação à cultura cartográfica. Entenda-se por cultura o costume de uso de mapas e seu manuseio tanto pra consultas rápidas como para análises mais complexas. Essa deficiência pode acarretar em subutilização do SIG, fazendo com que sua potencialidade seja pouco explorada e renegada a um segundo plano.

b. Qualidade dos dados

Uma característica importante dos SIG é o fato de serem bons na combinação de diferentes tipos de dados. Independente do fato dos dados se referirem a diferentes datas de levantamento, diferentes graus de resolução espacial (escalas), ou uma combinação destes fatores, o software sempre possibilita sua combinação na apresentação dos resultados. Se a qualidade dos dados for duvidosa o resultado apresentado pelos SIG também será. Esse tipo de problema inviabiliza o uso do sistema por transforma-lo numa possível fonte de erros fornecendo informações equivocadas para a tomada de decisão.

Os cartógrafos, na compilação de mapas, têm trabalhado com diferentes grupos de dados por séculos, e tem com isso angariado alguma experiência nas transformações que são necessárias para ordenar e combinar grupos de dados de diferentes: resoluções (escalas), sistemas de referencia geodésicos, sistemas de projeções cartográficas e dados de levantamentos. Estes profissionais têm desenvolvido transformações e procedimentos (tais

como a generalização) que tratam destas diferenças e possibilitam a real integração de dados. Concomitantemente, eles têm desenvolvido a documentação técnica que descrevem todos as características relevantes dos dados (metadados) necessários para uma integração apropriada. Além disso, também têm influenciado por décadas a padronização destes documentos técnicos que qualquer grupo de dados pode ser facilmente manipulado (Figura 4.14). Conseqüentemente grande parte das metodologias para a determinação da qualidade de dados potencialmente disponível para os usuários de SIG tem origem na cartografia.

A avaliação da qualidade dos dados pode ser definida, dentro deste contexto, como uma medida da adaptabilidade dos dados para uma aplicação específica. Portanto tendo em mente tal definição o trabalho do produtor de dados espaciais para os SIG deve prever o uso de seus dados em ambientes diversos. Estes dados por sua vez devem garantir ao usuário final as condições de uso que satisfaçam as necessidades das aplicações dos SIG.

2.3.2 Fontes de erros dos dados

A obtenção de dados digital espacial a partir de cartas pré-existentes ocasiona uma série de problemas de tratamento de dados. Quando se opta pela digitalização vetorial de uma carta convertida e georreferenciada para o meio digital em formato matricial podem ocorrer diversos tipos de erros. O Quadro 2.1 relaciona as principais fontes de erros que podem ocorrer quando se considera o uso de dados desta natureza em SIG.

O quadro mostra três principais grupos de fatores que regem os erros que podem interferir no processamento de informações geográficas.

O grupo I inclui fontes de erros mais óbvias que são de fácil checagem; O grupo II contém fontes de erro mais subjetivas que podem geralmente ser detectadas quando se trabalha mais intimamente com os dados; O grupo II é o mais importante, pois inclui enganos, erros e descuidos que podem surgir como resultado da realização de certos tipos de processamento. O grupo III abrange os erros mais difíceis de serem acusados porque eles requerem um intimo conhecimento não somente dos dados, mas também da estrutura dos dados e o algoritmo usado.

As fontes de erro apresentadas servem como alerta aos usuários dos SIG, pois não é difícil encontrar problemas relacionados aos dados geográficos nestes sistemas.

Embora seja importante o conhecimento dos grupos de erros I e III, é o grupo II que causa mais incomodo aos produtores de dados geográficos.

Os erros oriundos da exatidão posicional das feições (número 9 da lista contida no quadro) acarretam problemas de ligação entre espaços geográficos contíguos e de posicionamento de feições a serem inseridas ou atualizadas. Estes erros em particular têm origem na realização do desenho cartográfico ou no processo de conversão deste desenho para o meio digital. Sendo neste último acarretado por deficiências na fase de registro do material cartográfico durante o processo de digitalização vetorial.

I. Fontes de erros óbvias
1. Idade dos dados
2. Cobetura da área – parcial ou completa
3. Escala do mapa
4. Densidade de observações
5. Relevância
6. Formato
7. Acessibilidade
8. Custo
II. Erros resultante de variações naturais ou de medições originais
1. Acurácia posicional
2. Acurácia do conteúdo – quantitativa e qualitativa
3. Fontes de variação dos dados:
- entrada de dados ou saída defeituosa
- ruídos de observação
- variação natural
III. Erros surgidos no processamento
1. Erros numéricos no computador:
- limitação da representação de números do computador
2. Enganos surgidos por meio de análises topológicas:
- emprego errado da lógica
- problemas associados com a sobre-posição de mapas
3. Problemas de classificação e generalização:
- metodologia
- definição do intervalo de classes
- interpolação

Quadro 2.1 - Fontes de possíveis erros para os SIG
Fonte: adaptado de Burrough, 1986.

Os erros devido à exatidão de conteúdo (número 2 do item II do quadro) dizem respeito à correção ou ausência de ruído nos atributos vinculados aos pontos, linhas e áreas contidos no BD geográfico. A exatidão qualitativa diz respeito a correção da definição dos elementos espaciais armazenados no BD. Por exemplo, um polígono pode estar

erradamente codificado como uma área de cultura ao invés de mata. A exatidão quantitativa por sua vez diz respeito ao nível de ruído que um valor dado pode assumir se comparado com os demais. Por exemplo, quando o uma cota apresenta um valor muito alto ao ser comparada com a altitude média de uma determinada região.

Os erros devido às fontes de variação nos dados (número 3 do item II do quadro) são provenientes enganos na entrada de dados num SIG. A princípio esses erros podem ser corrigidos facilmente quando se apresentam de maneira grosseira ou sistemática, entretanto, quando aparecem de forma aleatória e caracterizam deficiência de interpretação ou são introduzidos acidentalmente podem se tornar um problema de difícil solução. Quando se dá a entrada de dados geográficos diversos erros podem surgir. Em se tratando especificamente de dados geográficos de cunho espacial os erros devem ser rigorosamente tratados, pois deles dependem o bom funcionamento dos SIG.

Sendo assim, portanto, serão abordados neste trabalho de dissertação, somente os erros provenientes do grupo II.

2.3.3 Erros associados à digitalização vetorial de uma carta

Pode-se assumir implicitamente para estimação dos erros, conforme estabelecido pelo método Suíço [Bourrough, 1986], que o mapa “verdadeiro” tem unidades de mapeamento uniformes e fronteiras infinitamente finas. Na prática, entretanto, os melhores desenhos de mapas não são perfeitos e erros extras são introduzidos pelo processo de digitalização. Considerando que uma carta pode ser digitalizada no formato vetorial a partir do original matricial fica fácil visualizar o problema do exato posicionamento da linha que representará o desenho final do documento cartográfico naquele formato. Neste caso há duas fontes potenciais de erros: erros associados com a fonte da carta e erros associados com a representação digital da mesma.

Erros associados à origem dos dados

Apesar dos erros potencialmente corrigíveis do estiramento do papel e distorções nos mapas impressos ou no documento fonte surge erros devido à locação de limites simplesmente porque os desenhos dos limites não são infinitamente finos. O traçado de uma linha com 1mm de espessura sobre um mapa na escala 1:250.000 corresponde a

uma linha com largura de 250m; o mesmo traçado sobre uma escala 1:100.000, por sua vez, corresponde a uma largura de 100m. Dentro deste contexto se representarmos uma área sobre um mapa na escala 1:250.000 com dimensões 20 x 20mm teríamos uma área equivalente a 10% do desenho da área ocasionada pela espessura do traçado (Figura 2.23).

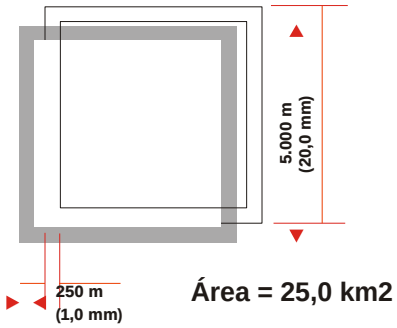


Figura 2.23 - Erro em área para a representação de um quadrado de 20 x 20 mm numa escala 1:250.000.

Hoje o desenho realizado em sistemas computacionais especialistas, como é o caso do CAD, permite se realizar o traçado das feições tomando por base vetores que possuem linhas infinitamente estreitas. Neste caso o senso comum sugere que a linha de divisão verdadeira seja tomada como um ponto médio da linha do desenho da carta. Esta seria a situação perfeita, entretanto o que ocorre na prática é que devido às condições impostas no processo de digitalização da linha existirá sempre uma área de incerteza sobre a qual o desenho final estará sujeito.

É o que ocorre quando as linhas existentes em originais cartográficos ou cartas impressas são convertidas por digitalização vetorial, pois o operador encarregado da digitalização nem sempre seguirá exatamente a metade da linha. Uma tentativa de contornar tal problema e minimizar este efeito é usar técnicas de vetorização baseadas em algoritmos que trabalham com estruturas de arquivos matriciais de 1bit (assunto já abordado).

Erros associados à representação digital dos dados

A representação de formas curvas depende do número de vértices usados para construí-las [Aldred, 1972]. Conseqüentemente, o erro relativo da digitalização vetorial de linhas retas é menor que o de digitalização de curvas complexas. Converter uma linha curva contínua de um mapa para um formato vetorial envolve um processo de amostragem, em

outras palavras, somente uma quantidade muito pequena dos infinitos pontos que descrevem a curva é capturada durante a digitalização vetorial conforme ilustrado na Figura 2.24 [Smedley & Aldred, 1980].

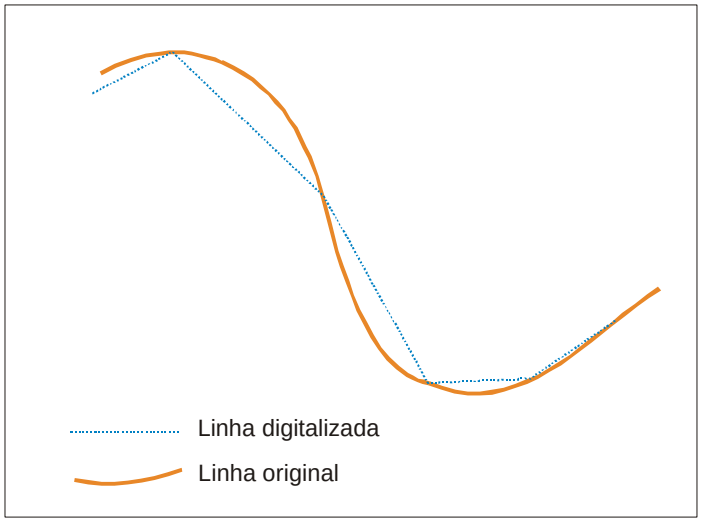


Figura 2.24 - Digitalização vetorial de uma linha curva.
Fonte: adaptado de Burrough, 1986.

Logicamente, limites de áreas nos mapas não devem ser entendidos como absolutos, mas sim contendo uma faixa de erro associado ou intervalo de confiança. MacDougal (1975) sugere que a inexatidão do limite pode ser estimada por

$$H = \sum_{i=1}^{N} h_i l_i / T$$

Onde,

h_i é o erro horizontal (em desvio padrão) da linha;

l_i é o comprimento;

N é o número de linhas limite no mapa;

T é a área total do mapa.

Se todas as linhas limites são do mesmo tipo (por exemplo, todas representam limites de tipo de solos) a equação apresentada anteriormente é simplificada para a forma:

$$H = h L / T$$

Perkal (1966) definiu um “epsilon”, representado por uma área de incerteza ao longo de uma linha cartográfica, como uma maneira de generaliza-la objetivamente. O conceito pode ser usado eficientemente para indicar a região de influência ao longo da linha onde podem ocorrer os erros de digitalização vetorial. Blackmore (1984) discute a aplicação deste conceito com respeito à aplicação do bem conhecido problema “ponto-em-polígono”. Ele mostra que a questão “Onde o ponto xy se encontra em relação ao polígono A?” resulta em cinco possíveis respostas:

(1) definitivamente contido em A;

(2) definitivamente não contido em A;

(3) possivelmente contido em A;

(4) possivelmente não contido em A;

(5) sobre a linha limite de A digitalizada em forma de vetor.

Estas situações são ilustradas na Figura 2.25. A situação na qual o ponto pode ser considerado “definitivamente contido em A” refere à posição da área delimitada dentro da faixa de incerteza; “possivelmente contido em A” esta relacionada à posição do ponto que cai dentro da faixa de incerteza para dentro do polígono. “Possivelmente não contido em A” é relacionado a um ponto que cai dentro da faixa de incerteza, mas fora do polígono, tecnicamente falando o ponto neste caso deveria ser considerado como se caísse fora do polígono, mas ele poderia de certa forma estar dentro do polígono “verdadeiro” se tivesse sido erroneamente digitalizado. A condição de um ponto estar exatamente posicionado sobre a linha digitalizada delimitadora do polígono é um caso raro, porém pode ocorrer.

Este tipo de análise minuciosa ajuda a visualizar e compreender melhor a natureza dos erros gerados no posicionamento de vértices durante processo de vetorização. Atualmente dentro do âmbito da DSG uma linha formada por um vetor é considerada posicionada corretamente se a mesma se encontra no meio da linha matricial gerada no processo de digitalização. A obediência deste procedimento na vetorização, entretanto nem sempre é possível devido a fatores técnicos que podem ser mais bem compreendidos quando se trabalha com processo de vetorização manual em tela ou em mesa digitalizadora (Capítulo 2 – Parte II) e, portanto é aceito uma certa tolerância quanto ao posicionamento dos vértices dos vetores em relação à linha matricial. Segundo o Manual de Vetorização da

DSG para o traçado das linhas são toleráveis erros máximos de posicionamento do vetor em relação à linha matricial em até 0,2 mm na escala da carta.

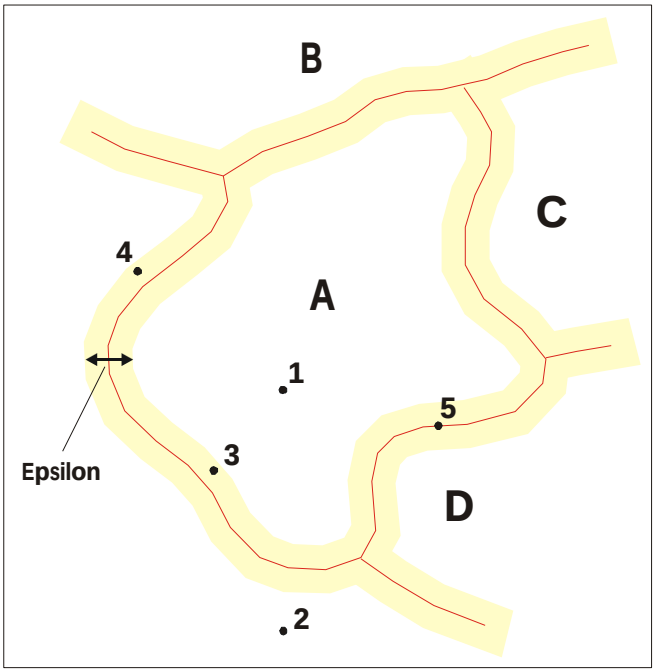


Figura 2.25 - “Onde o ponto xy se encontra em relação ao polígono A?”: (1) definitivamente contido em A; (2) definitivamente não contido em A; (3) possivelmente contido em A; (4) possivelmente não contido em A; (5) sobre a linha limite de A digitalizada em forma de vetor.
Fonte: adaptado de Burrough, 1986.

Erros associados à sobreposição de camadas de elementos de área

A associação espacial de camadas de informações contendo elementos de área é uma prática comum quando se trabalha com dados espaciais para fins cartográficos ou em análises num SIG. A coincidência entre os vértices dos vetores formadores dos limites dos elementos de área é uma condição desejável nestas operações.

Antes do advento dos mapas digitais em formato vetorial, o processo de confecção cartográfica era executado por meio do uso de bases transparentes e os limites coincidentes das áreas eram estabelecidos por meio do uso linhas criadas pelo traçado realizado por canetas de pontas relativamente grossas. A sobreposição de áreas contidas nas cartas em ambiente digital prometia melhores resultados porque todos os limites eram supostamente codificados com precisão, mas de fato a nova tecnologia trouxe um dos problemas mais difíceis e mais pesquisados na cartografia computacional.

Devido aos erros associados com a representação digital da carta (item anterior), os polígonos geralmente destoam quando camadas de informações distintas são sobrepostas. Isto é facilmente percebido quando consideramos a sobreposição de uma camada contendo o traçado de uma rede hidrográfica com uma camada contendo informações de cobertura vegetal. Se as informações contidas nas duas camadas foram vetorizadas individualmente certamente os elementos de área contíguos no espaço da carta não irão coincidir. Isto exige um tratamento adicional do traçado das feições vetorizadas para que as mesmas não apresentem inconsistências topológicas graves.

Basicamente a solução adotada para solucionar tal tipo de problema é o ajuste das fronteiras dos elementos de área das camadas envolvidas. O ajuste ideal seria usar a área de incerteza existente no limite de cada polígono para definir a fronteiras mais provável entre as áreas, entretanto esse procedimento é muito complexo e envolveria fatores que nem sempre poderiam ser determinados.

Uma solução mais prática é adotar os limites existentes em uma camada como referência para definição os limites coincidentes da camada a ser sobreposta. A questão que pode ser levantada neste ponto é qual camada deve ser tomada como referência, para responder esta questão, ou melhor, para orientar esta decisão pode se levar em conta à experiência e a lógica dos procedimentos de aquisição de dados. A prática mostra que existem feições de área que são mais bem capturadas que outras no processo de aquisição de dados. Por exemplo, o traçado de rios de margem dupla, lagos e outras massas d'água são mais bem definidos que áreas de vegetação, pois estas muitas vezes ou apresentam falta de padrões bem definidos para serem caracterizadas ou apresentam uma faixa de transição não muito nítida entre as diversas espécies vegetais que compõe estes tipos de feições.

Erros comuns ocasionados durante a aquisição de dados

No processo de aquisição de dados vetoriais ocorrem erros potenciais que devem ser eliminados para que se possa garantir o relacionamento topológico entre as feições dentro do ambiente do SIG. A Figura 2.26 ilustra os erro mais comuns que ocorrem durante o processo de vetorização. Os erros cometidos durante o processo de digitalização são basicamente seis:

Vértice após o ponto (*overshot*)

Este tipo de engano ocorre quando o operador posiciona o vértice de uma linha vetorial após a posição desejada. Geralmente o erro pode ser evitado por meio do uso de ferramenta existente no CAD que permitem ajustar a posição final da linha de acordo com a posição de um outro vértice ou exatamente sobre o traçado de uma outra linha vetorial (*snap*). Quando por algum motivo não é possível utilizar-se este tipo de ferramenta o posicionamento além do ponto desejado pode ser corrigido por meio de ferramentas de edição do CAD.

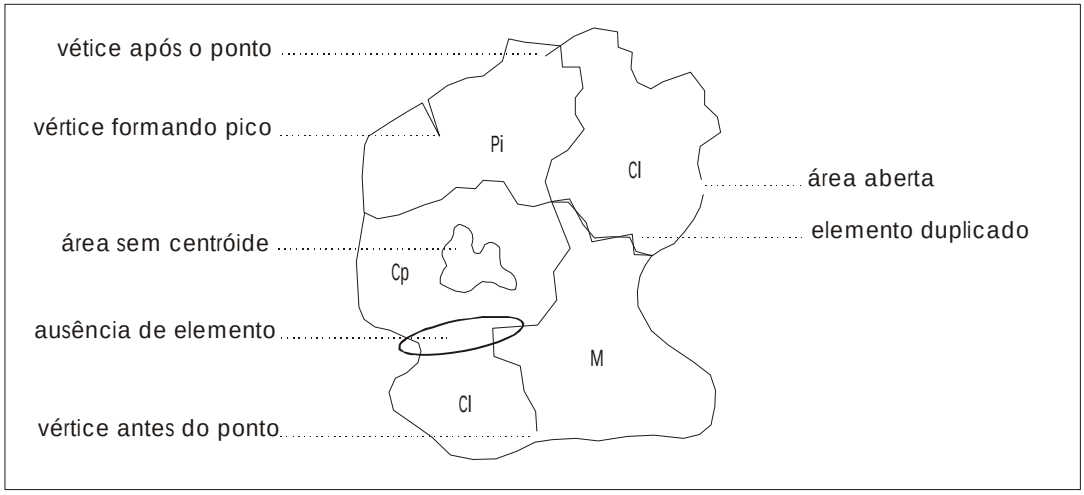


Figura 2.26 - Erros potenciais do processo de vetorização.
Fonte: adaptado de Kraak & Ormeling, 1996.

Vértice formando pico

Os vértices em forma de pico são caracterizados por pontos posicionados fora do padrão de concordância de uma reta ou uma curva. São ocasionados por erro do operador que por algum motivo qualquer possa ter deixado de percorrer adequadamente um trecho de uma feição conforme o desenho do mapa. A correção deste tipo de erro implica em eliminar o vértice em questão sem interromper a continuidade da linha vetorial. Isto pode ser realizado por meio de ferramenta de edição num CAD.

Área sem centróide

O centróide é o elemento responsável pela definição do que representa uma determinada área ou polígono fechado. Este elemento auxiliar pela definição de áreas no

formato vetorial é extremamente importante, pois por meio dele é possível se identificar exatamente o que representa as diversas áreas existentes em uma carta. A ausência deste tipo de indicador pode causar uma não classificação da área e conseqüente erro de interpretação de dados, sua correção é realizada introduzindo-o após ter sido detectado sua falta no processo de revisão.

Ausência de elemento

A ausência de elemento é caracterizada pela falta de feições que deixaram de ser capturadas para dentro do arquivo final de vetorização. Geralmente é causado por falta de atenção do operador ou por limitações de tamanho de arquivo do software utilizado, por exemplo, as primeiras versões do CAD *MicroStation* permitiam trabalhar somente com um tamanho de arquivo equivalente a 32Mb, quando se passava deste tamanho de arquivo ocorria uma eliminação de elementos caso se desejasse inserir mais feições (vértices). A correção deste tipo de falta é a digitalização vetorial novamente da feição após ter sido detectada numa revisão a falta do elemento.

Vértice antes do ponto (*undershot*)

Similar ao vértice após o ponto, só que neste caso o operador posiciona o vértice antes do ponto desejado.

Área aberta

Erro caracterizado pela falta de junção entre dois vértices consecutivos de uma linha de contorno delimitadora de uma área. O erro é causado durante a operação de vetorização por distração do operador ou por falha do processo de vetorização utilizado. A correção de tal falta é realizada por meio da captura do elemento faltante a partir do original usado.

Elemento duplicado

Ocorrência de elementos duplicados são considerados erros, pois dentro de um BD geográfico é necessário garantir a unicidade de dados, ou seja, cada dado armazenado deve ser único. Este tipo de erro pode ser cometido durante o processo de vetorização

quando se tem de digitalizar com uma linha áreas contíguas. A correção do erro implica dna eliminação de um dos elementos, e geralmente é realizada após a constatação de duplicidade de linha por meio do uso de ferramentas de seleção ou intercalação de camadas no ambiente de CAD.

Tendo em vista que os erros apresentados impossibilitam a aplicação de regras topológicas, esses devem ser corrigidos para que as feições vetorizadas sejam levadas para dentro de um ambiente de SIG. Os procedimentos adotados para correção dos erros foram reunidos numa fase de produção denominada por validação, neste caso em particular validação geométrica.

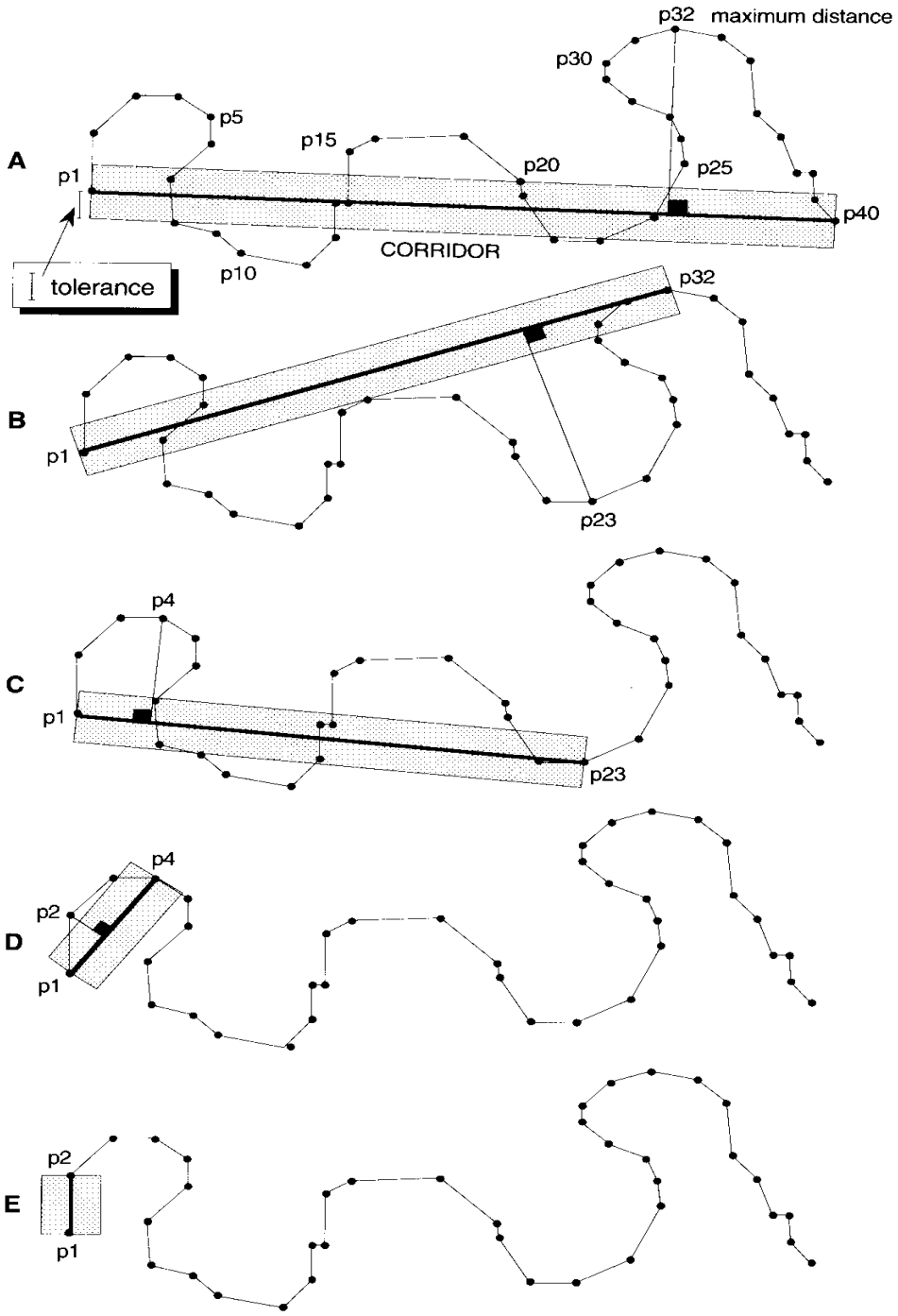
Todos os erros apresentados são passiveis de identificação por meio de inspeção visual e sua correção é possível por meio da iteração do operador com as ferramentas de CAD existentes no ambiente de trabalho. Entretanto, devido à natureza de alguns erros é possível gerar rotinas computacionais baseadas em algoritmos geométricos para a sua detecção e correção. A validação geométrica adotada pelas instituições responsáveis pelo mapeamento sistemático adota este princípio para desenvolvimento de rotinas de automatização de processos.

Basicamente, os erros geométricos apresentados que podem ser corrigidos via emprego de rotinas computacionais são: vértice após o ponto, vértice antes do ponto, vértice formando pico, área aberta e elemento duplicado. Os demais erros, a saber, ausência de elemento e falta de centróide devem ser corrigidos “manualmente” pelo operador após terem sido detectados pela revisão. O Capítulo 3 tece alguns comentários em relação à validação de BD geográficos e descreve pormenorizadamente o processo de validação de arquivos vetoriais atualmente realizado no âmbito da DSG.

2.3.4 Algoritmo Douglas-Peucker

Trata-se do método mais utilizado pelos sistemas de informações geográficas. Concebido inicialmente para resolver o problema do número excessivo de pontos resultantes da conversão de dados gráficos para o formato digital, o método de Douglas-Peucker baseia-se na seguinte idéia: se nenhum ponto da linha encontra-se mais afastado do que uma certa distância vertical ao segmento de reta que liga os extremos da linha, então esse segmento de reta é suficiente para representar a linha. Este método é considerado uma

técnica global de generalização, pois analisa cada linha como um todo. A Figura 2.27 ilustra a aplicação do algoritmo de Douglas-Peucker.



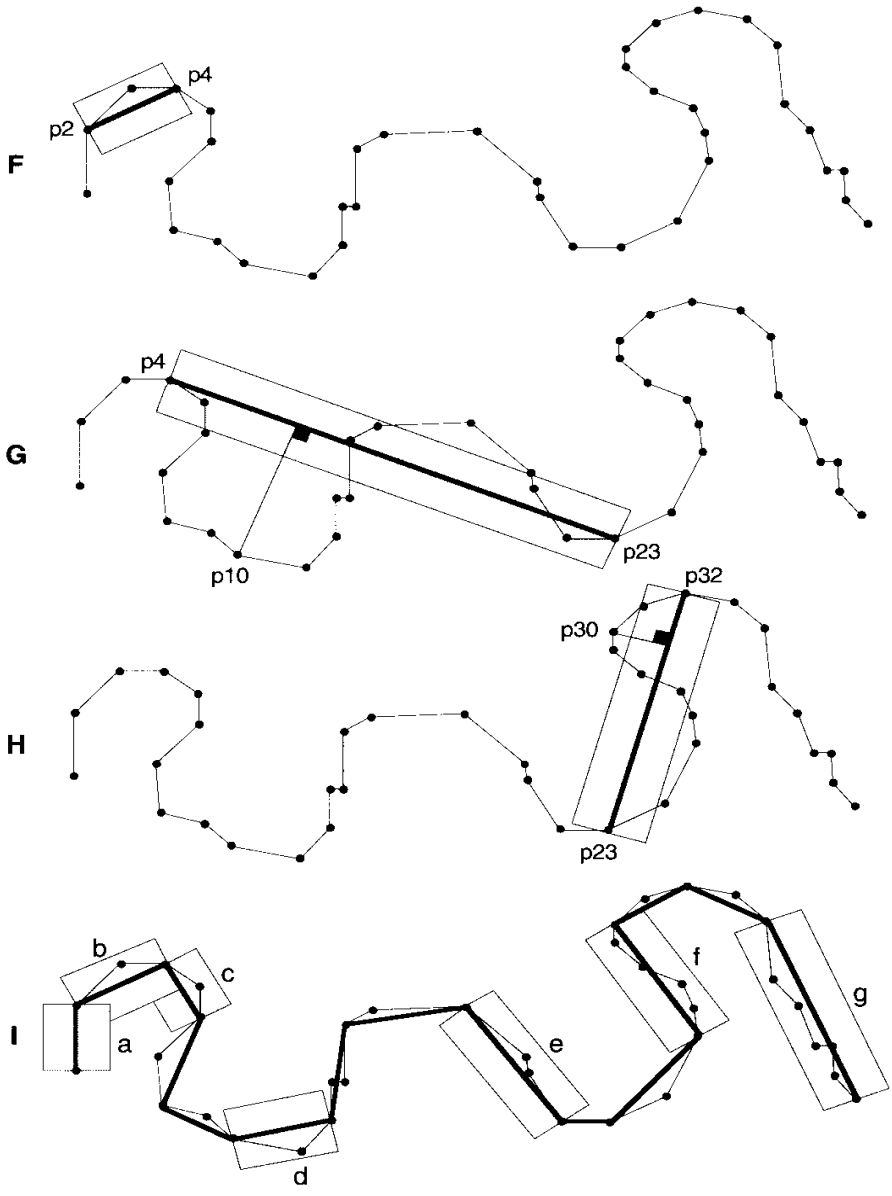


Figura 2.27 - Lógica do algoritmo Douglas Peucker, para eliminação de vértices em excesso.
Fonte: Guia do Usuário do Spring, 2005.

Parte IV – O AMBIENTE DE BANCO DE DADOS ORIENTADO A OBJETO

2.4 Banco de Dados Orientado a Objeto

Atualmente a técnica da orientação a objetos tem sido usada para projetar e implementar sistemas de natureza variada. Com relação a bancos de dados, essa técnica tem sido empregada, com predominância, nos casos aonde os dados envolvidos na aplicação considerada apresentam estrutura complexa [Vieira M, 2001]

Basicamente as diferenças existentes entre os modelos de dados tradicionais (relacional, hierárquico e em redes) e os modelos de dados orientados a objetos residem na maneira como eles enxergam os dados.

Os modelos de dados tradicionais vêem os dados como uma coleção de tipos de registros ou relações, cada um tendo uma coleção de registros ou tuplas armazenadas em tabelas. Já num modelo de dados orientado a objetos um banco de dados é considerado como uma coleção de objetos do mundo real.

Embora a informação sobre objetos complexos do mundo real possa ser espalhada em tabelas relacionais, a meta dos bancos de dados orientados a objetos é manter uma correspondência direta entre os objetos do mundo real e os do banco de dados, podendo estes ser identificados e manipulados como um todo. Representar um objeto complexo no modelo relacional significa que o objeto tem que ser subdividido em um grande número de tuplas, o que leva à necessidade de realizar um considerável número de operações de junção para recuperar o objeto.[Vieira M, 2001]

Os conceitos oriundos da orientação a objetos permitem a criação de uma base sólida para aplicações de banco de dados mais avançadas, como por exemplo, aplicações de engenharia tais como: CAD/CAM (*Computer Aided Design/Computer Aided Manufacturation*), CASE (*Computer Aided Software Engineering*) e Sistemas de Informação Geográfica.

Essas aplicações têm requisitos e características que diferem das aplicações comerciais tradicionais, tais como estruturas de dados mais complexas, transações de duração mais longa, novos tipos de dados para armazenar imagens ou textos longos e a necessidade de definição de operações não padrões específicas da aplicação. Os bancos de dados orientados a objetos foram propostos para dar suporte às necessidades dessas aplicações mais complexas.

Os modelos de dados orientados a objetos usam os conceitos de abstração de dados dos modelos semânticos (classificação, generalização e agregação) e incorporam outros conceitos.

2.4.1. Conceitos Básicos

Alguns conceitos encontrados nas linguagens de programação orientadas a objetos (LPOO) são também aplicados nos modelos de dados orientados a objetos, porém bancos de dados requerem alguns conceitos próprios. Os objetos, em uma LPOO, existem somente durante a execução do programa e são por isso chamados de **transitórios**. Um banco de dados orientado a objetos pode estender a existência dos objetos de modo que eles sejam armazenados permanentemente, isto é, os objetos são **persistentes** (persistem após o término do programa e podem ser recuperados posteriormente e compartilhados por outros programas).

A seguir são apresentados os principais conceitos envolvidos em bancos de dados orientados a objetos.

Objetos e Identidade

Cada entidade do mundo real é modelada como um objeto.

A cada objeto são associados um estado e um comportamento: o *estado* é representado pelos valores dos atributos do objeto; o *comportamento* é definido pelos métodos que agem sobre o estado do objeto pela invocação de operações.

A cada objeto armazenado no banco de dados é associado um identificador único (**OID -** *Object Identifier*), que é gerado pelo SGBDOO (sistema de gerenciamento de banco de dados orientado a objetos). O valor do OID não e visível ao usuário, mas é usado

internamente pelo sistema para identificar cada objeto de forma única e criar e gerenciar referências entre objetos.

A principal propriedade de um OID é que ele é imutável, isto é, o valor do OID de um particular objeto não deve mudar. O SGBDOO deve ter algum mecanismo para gerenciar OIDs e preservar a propriedade de imutabilidade. É também desejável que cada OID seja usado somente uma vez, isto é, os OIDs dos objetos que são removidos do banco de dados não são reaproveitados.

As duas propriedades acima implicam que o OID não deve depender de nenhum valor de atributo do objeto. Geralmente é considerado não apropriado basear o OID no endereço físico do objeto no meio de armazenamento, uma vez que o endereço físico pode mudar após a reorganização do banco de dados. Entretanto, alguns sistemas usam o endereço físico como OID, para aumentar a eficiência de recuperação do objeto. Nesse caso, se o endereço físico do objeto muda, pode ser colocado um ponteiro indireto no primeiro endereço, indicando a nova localização física do mesmo.

É mais comum usar inteiros longos como OIDs e uma função hash para mapear o valor do OID para o endereço físico do objeto.

Valores

A maioria dos SGBDOOs representam as entidades primitivas, tais como inteiros ou caracteres, por valores (não possuem OIDs), enquanto as entidades não primitivas são representadas como objetos. Já outros sistemas, como o OO, permitem a definição de valores complexos que não podem ser compartilhados por outros objetos, uma vez que valores não possuem OIDs.

Estrutura do objeto

O valor de cada atributo de um objeto pode ser:

- atômico: *integer*, *real*, *character*, *boolean*, etc.
- complexo: definido através de construtores: tuple, set, list, bag e array.

O construtor de tipo *tuple* serve para agregar informações afins. É freqüentemente chamado de tipo estruturado, pois corresponde ao construtor *struct* nas linguagens de programação C e C++.

Os construtores de tipo *set*, *list*, *array* e *bag* são chamados de tipos de coleção e servem para definir atributos multivalorados. Podem ser não ordenados (*set* e *bag*) ou ordenados (*list* e *array*). Em um set não pode haver dois elementos com o mesmo valor, enquanto que na *bag* isso é possível.

OIDs x Chaves Primárias

Nos modelos orientados a objetos não existe o conceito de chave primária como acontece no modelo relacional. Ao invés disso, existem os OIDs dos objetos que, como já dito, são criados e mantidos pelo SGBDOO e não são de acesso do usuário.

As vantagens do uso de OIDs com relação às chaves são:

- os programadores de aplicações não precisam se preocupar com a seleção de chaves para as várias classes de objetos;

- obtém-se melhor desempenho, pois os OIDs são implementados em baixo nível pelo sistema;

- embora as chaves sejam mais significativas ao usuário, muitas vezes o mesmo precisa usar códigos artificiais, sem significado semântico, para poder identificar as tuplas de uma relação.

Objetos Complexos

A composição estrutural de um objeto é definida através de um conjunto de atributos. O valor de cada atributo pode ser primitivo, um objeto ou uma combinação dos construtores tupla, lista, *array*, conjunto ou *bag*, envolvendo outros objetos ou não. Objetos complexos são definidos através de construtores envolvendo outros objetos.

Quando o valor de um atributo de um objeto O é um objeto O', o sistema armazena o identificador de O' em O ou todo o valor complexo é armazenado no atributo do objeto.

Encapsulamento

A cada objeto está associada sua estrutura e seu comportamento (os métodos ou operações). O comportamento é armazenado no banco de dados junto com a estrutura do objeto.

O conceito real de encapsulamento determina que somente as operações sobre os objetos são visíveis e sua estrutura é escondida. Em banco de dados a noção de invisibilidade da estrutura do objeto é afrouxada. É desejável, por exemplo, poder consultar os atributos do objeto através de uma linguagem de consulta. Assim, a maioria dos SGBDOOs permite acesso direto aos atributos fornecendo operações definidas pelo sistema para a leitura e modificação dos atributos, o que livra o usuário da incumbência de implementar uma considerável quantidade de métodos cujo único propósito é ler e escrever os valores dos vários atributos dos objetos. Isso é um exemplo de violação do encapsulamento permitida pelos SGBDOOs. Esses sistemas, porém, possuem mecanismos para que o usuário possa proteger o acesso aos atributos dos objetos, caso desejável. O sistema OO, por exemplo, permite o usuário estabelecer quais atributos e métodos são visíveis na interface do usuário, através da declaração *public*, o que permite serem invocados por qualquer outro objeto. Os não visíveis são referidos como *private*.

Métodos

Os objetos nos SGBDOOs são manipulados através de métodos. Em geral, a definição de um método consiste de *assinatura* e *corpo*. A assinatura especifica o nome do método, os nomes e classes dos argumentos e a classe do resultado, se existir. O corpo representa a implementação do método e consiste de um conjunto de instruções expressas em uma dada linguagem de programação.

Tipos e Classes

Um *tipo* modela as características comuns de um conjunto de objetos e corresponde à noção de tipos abstratos de dados. Uma *classe* é um conjunto de objetos que tem exatamente a mesma estrutura interna, isto é, os mesmos atributos e mesmos métodos.

Os modelos de dados orientados a objetos usam o conceito de classe como uma base para instanciação.

Herança

É um mecanismo de reuso muito poderoso. Com a herança, uma *classe* considerada *superclasse* pode definir uma *subclasse*. A subclasse herda os atributos,

métodos e mensagens de sua superclasse e pode ter atributos específicos, métodos e mensagens adicionais.

Exemplo: Considere duas classes com informações sobre um conjunto de ônibus e caminhões. As características das duas classes são mostradas na Figura 2.28 cuja notação gráfica utilizada representa cada classe por um retângulo dividido em 3 partes. A parte superior contém o nome da classe; a do meio contém os atributos e a inferior contém os métodos definidos pelo usuário. Como as duas classes possuem algumas características em comum, pode-se criar a classe Veículo para conter essas características, como na Figura 2.29. Somente as características próprias de cada subclasse são mantidas na mesma.

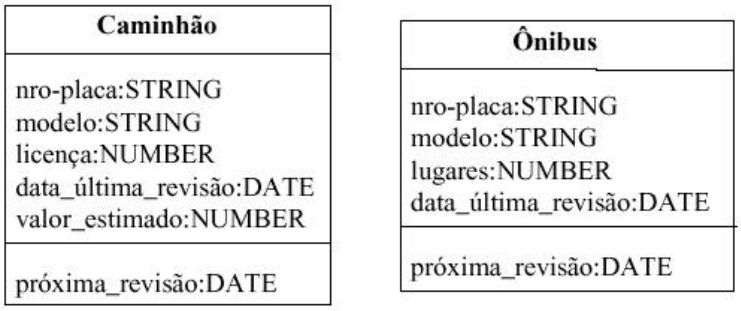


Figura 2.28 - Características de duas classes representadas graficamente por retângulo dividido em três partes: a superior contém o nome da classe; a do meio contém os atributos e a inferior contém os métodos definidos pelo usuário.
Fonte: Vieira , 2001.

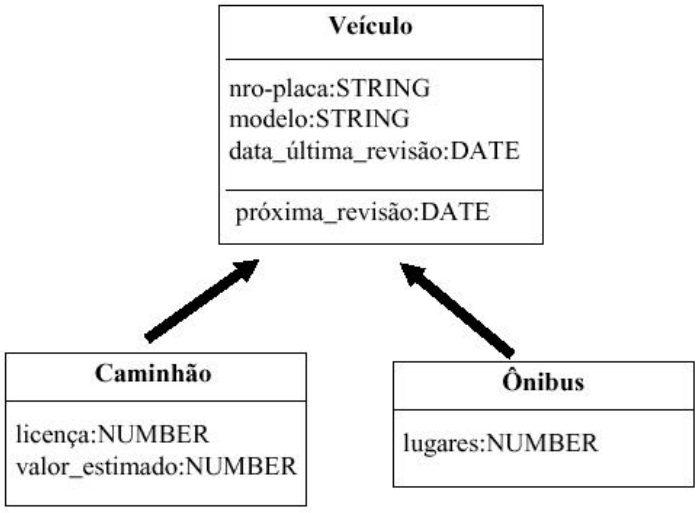


Figura 2.29 - Hierarquia e herança. Criação da classe Veículo para conter as classes Caminhão e Ônibus. Somente as características próprias de cada subclasse são mantidas na mesma.
Fonte: Vieira , 2001.

Vantagens da utilização de hierarquias de classe:

- ? diminui a quantidade de código a ser escrito; e,
- ? propicia uma descrição mais precisa e concisa da realidade.

Em certos sistemas, uma classe pode ter várias superclasses, em cujo caso diz-se que ocorre “herança múltipla” (Figura 2.30), enquanto outros impõem a restrição de uma única superclasse, dita “herança simples”.

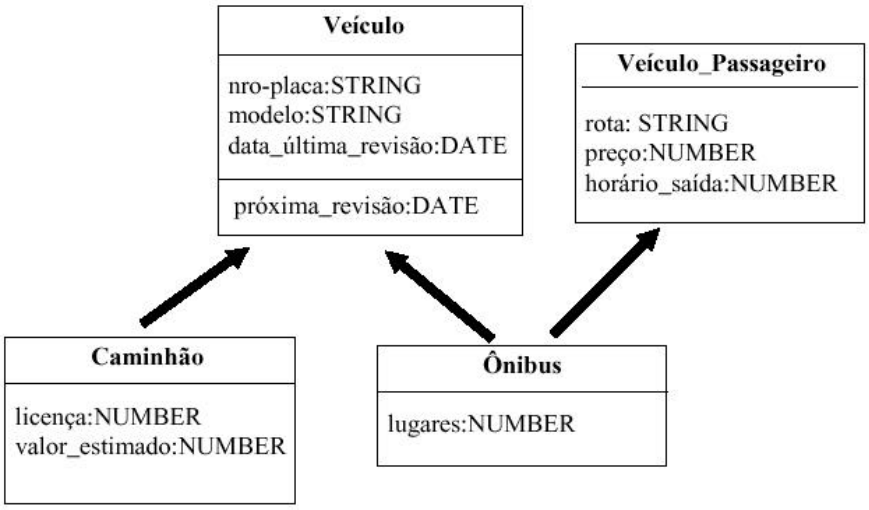


Figura 2.30 - Exemplo de herança múltipla para ao caso da classe Ônibus e herança simples para a classe Caminhão.
Fonte: Vieira, 2001.

A herança múltipla pode provocar problemas de conflitos, como por exemplo, duas ou mais superclasses podem ter um atributo com o mesmo nome, mas com diferentes domínios. Esses conflitos precisam ser tratados pelo sistema. Se existe uma relação de inclusão entre os domínios, então o domínio mais específico será escolhido como o domínio para a subclasse. Por exemplo, se na classe Veículo existir o atributo combustível cujo domínio é: ‘gasolina, álcool, diesel’ e em Veículo_Passageiro existir também o atributo combustível cujo domínio é ‘diesel’, a classe Ônibus herdará o atributo combustível cujo domínio será ‘diesel’ (Figura 2.31), isto é, o domínio mais restrito. Se essa relação não existe, uma solução adotada é a escolha do domínio com base na ordem de precedência entre as superclasses. Outros sistemas deixam por conta do usuário a resolução do conflito.

Em um esquema de banco de dados, as classes podem ser organizadas em uma hierarquia de herança, formando um grafo acíclico dirigido.

Polimorfismo

Os SGBDOOs oferecem o recurso de polimorfismo de operações, também conhecido como sobrecarga de operador (*overloading*). Outros conceitos relacionados com o polimorfismo são os de ligação tardia (*late binding*) e redefinição de operação (*overriding*).

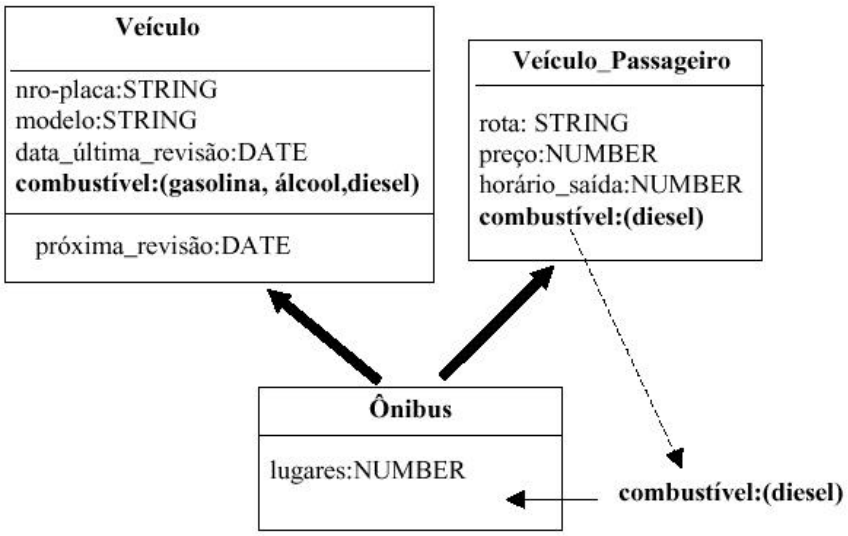


Figura 2.31 - Herança múltipla.
Fonte: Vieira, 2001.

Para melhor expor esses conceitos, considere uma operação *display* que recebe um objeto como entrada e apresenta o objeto na tela. Se o objeto for:

- uma imagem: deseja-se apresentar a imagem;
- uma pessoa: deseja-se apresentar os dados sobre a pessoa (nome, endereço, etc); e,
- um gráfico: deseja-se apresentar uma representação gráfica.

Usando um sistema convencional, seriam necessárias três operações: display_pessoa, display_figura e display_gráfico, como mostrado no Quadro 2.2.

Em um sistema orientado a objetos, a operação *display* pode ser definida em uma classe mais geral. A operação tem um único nome e pode ser chamada

indiscriminadamente por vários objetos. A implementação da operação é *redefinida* para cada uma das subclasses. Essa redefinição é chamada *overriding*.

```
for x in X do
begin
    case of type(x)
    pessoa: display_pessoa(x);
    figura: display_figura(x);
    grafico: display_grafico(x);
    end;
end;
```

Quadro 2.2 - Código convencional para a execução de operação display
que recebe um objeto como entrada e o apresenta na tela

O sistema decide qual implementação usar para execução. Assim, o código dado é simplificado para:

```
for x in X do display(x)
```

A ligação do nome da operação com a implementação correspondente se realiza no tempo de execução. Essa ligação retardada é denominada *late binding*.

A sobrecarga (*overloading*) de operador refere-se ao uso do mesmo símbolo de operador para denotar operações distintas sobre diferentes tipos de dados.

2.4.2 Modelagem de dados orientada a objetos para fins cartográficos

A motivação mais imediata para a adoção dos princípios da orientação a objetos para a modelagem de dados geográficos é que o mundo real pode ser visto como uma variedade de objetos inter-relacionados.

Os objetos podem ser vistos em diferentes níveis de detalhe. Por exemplo, para um observador, uma área urbana pode ser vista como um objeto indivisível, sem levar em consideração sua composição em áreas residenciais, industriais ou comerciais. Outro observador pode estar interessado em analisar com mais detalhe, por exemplo, uma área urbana residencial considerando a divisão das quadras, dos lotes, as edificações residenciais

e o poder aquisitivo de cada unidade residencial. Neste caso, as edificações residenciais são consideradas objetos que fazem parte da composição do objeto área urbana.

Essa maneira natural de ver a composição dos objetos deve ser conservada na modelagem dos objetos e é o objetivo dos modelos de dados orientados a objetos, fornecendo uma representação mais natural do mundo real.

Esta visão de modelagem de dados aplicada nos SIG atual é uma realidade para o qual os produtores de dados geográficos não podiam ficar alheios. A modelagem no ambiente orientado a objeto trás diversas vantagens para fins de automatização de processos produtivos dentro cartografia digital.

Assim como na modelagem de dados convencional, a modelagem de dados orientada a objetos é realizada em 2 fases:

1. Projeto conceitual:

Essa fase visa o projeto de um esquema conceitual que apresente uma abstração do problema do mundo real. Uma diferença quando se trata de projeto conceitual do banco de dados orientado a objetos é que, além da definição da estrutura dos objetos, também são definidos os métodos que manipulam esses objetos. Assim, toda a funcionalidade do sistema é definida juntamente com a estrutura dos objetos.

2. Projeto lógico:

É o projeto de uma estrutura lógica, representando o esquema lógico do banco de dados orientado a objetos, com base no esquema conceitual.

Dentro da cartografia a modelagem de dados do espaço mapeado é um fator condicionante não só para a aplicação final dos SIG como também para o uso destes sistemas na automatização de processos produtivos. Portanto dentro deste enfoque a DSG procurou desenvolver a modelagem do espaço territorial do Brasil de forma a contemplar todas as feições encontradas nas escalas do mapeamento sistemático.

Modelagem de dados da DSG

A necessidade de dispor de um Banco de Dados Geográficos capaz de suportar a automatização da cartografia e armazenar os dados obtidos no mapeamento sistemático

disponibilizando pesquisas do Espaço Geográfico Brasileiro (EGB) fez com que a DSG estuda-se o desenvolvimento de uma modelagem conceitual, com mecanismos que garantissem a integridade dos dados produzidos. Isto foi possível somente a pós a definição de um padrão estruturado para intercâmbio dos dados digitais, por meio da TBCD e dos procedimentos de validação e estruturação em ambiente de CAD.

A modelagem do Espaço Geográfico Brasileiro foi elaborada durante o período de 1999 a 2001 na 1ª Divisão de Levantamento (1ªDL). A versão final da modelagem estabelece a existência de oito Temas (Figura 2.32) aos quais são diretamente vinculados as categorias da TBCD. A estrutura básica que compõe cada Tema por sua vez pode ser visualizada na Figura 2.33.

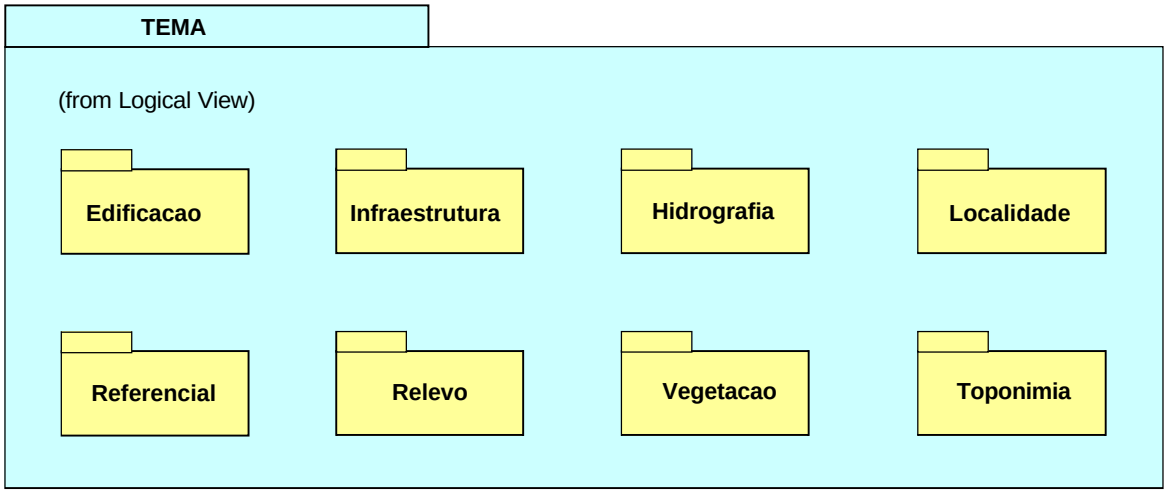


Figura 2.32 - Temas da modelagem do Espaço Geográfico Brasileiro.
Fonte: DSG, 2003.

A implementação da modelagem em ambiente orientado a objetos, entretanto foi possível somente após a adoção do sistema Gothic em 1999. A tarefa de implantar a modelagem as atividades de produção cartográfica ficou a encargo da 1ª Divisão de Levantamento (1ªDL) fazendo com que esta unidade se torna-se pioneira no uso de tal sistema no âmbito de Brasil.

2.4.3 Sistema *Gothic*

O banco de dados *Gothic* é produzido pela empresa *Laser-Scan Limited*, situada em Cambridge, Inglaterra, e foi projetado para construção de Sistemas de Informação que processam e usam dados geográficos. Assim o ambiente *Gothic* é um conjunto de

ferramentas para construir, e processar aplicações para informação relacionadas com o espaço. Estas aplicações capacitam a captura, armazenamento, pesquisa e a manipulação de dados espaciais. O conjunto de ferramenta é projetado para vários *hardwares* e configurações de sistemas operacionais. O sistema original foi desenvolvido para UNIX, porém pode ser operado em uma plataforma Windows NT4.0 ou superior.

O ambiente pode ser diferente de um usuário para outro. Isto é devido aos diferentes módulos de funcionalidades, da plataforma de *hardware* e do sistema operacional instalados. Porém, há uma estrutura de *software* essencial na qual as várias opções do produto se ajustarão.

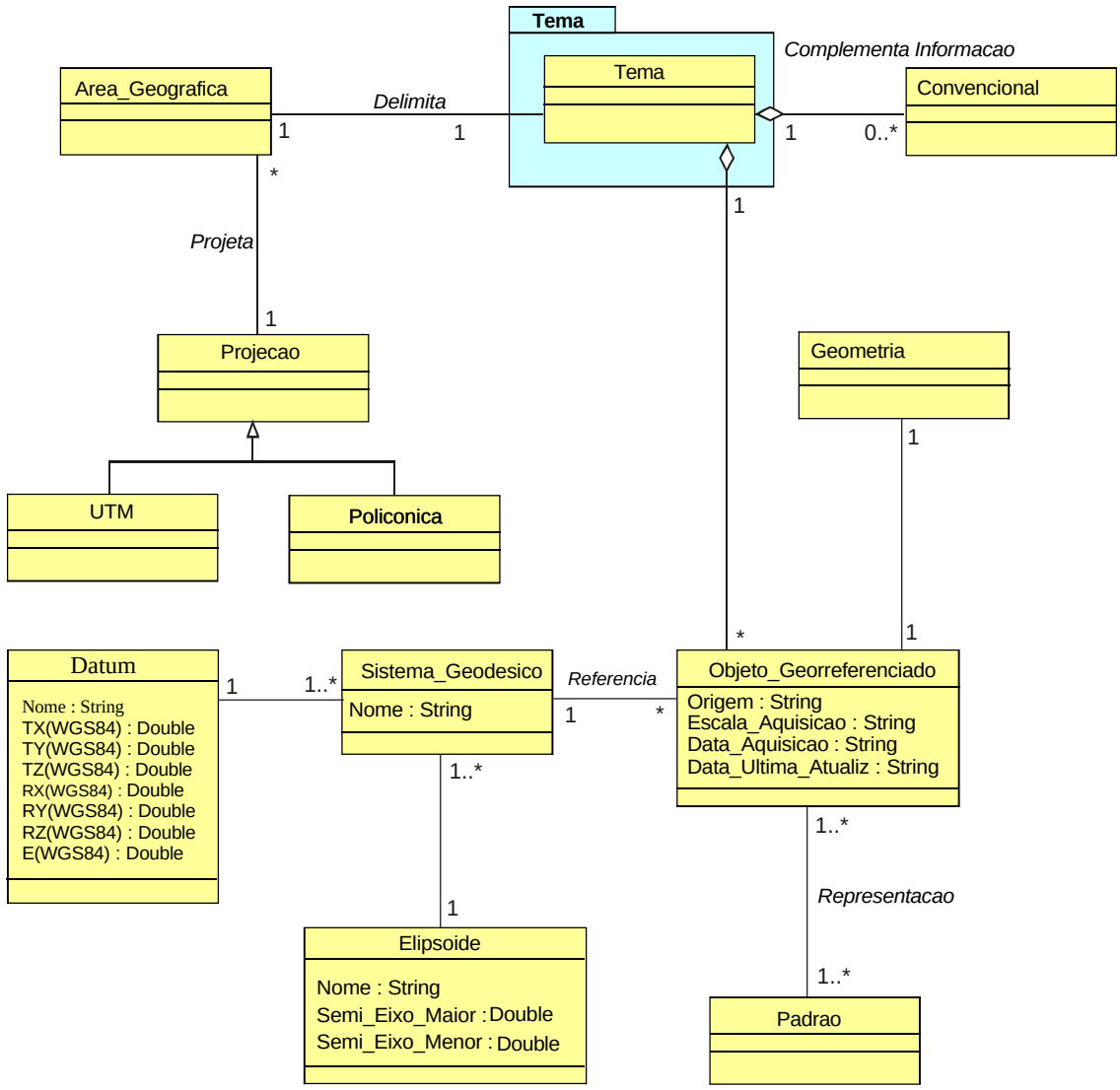


Figura 2.33 - Estrutura básica dos Temas da modelagem do espaço Geográfico Brasileiro.
Fonte: DSG, 2003.

2.4.3.1 Arquitetura do sistema

A OGC impõe aos seus consorciados um Modelo Abstrato que, entre vários objetivos, garante as funcionalidades essenciais para um banco de dados geográficos. Na Figura 2.34, são apresentados os principais componentes do *Gothic*, que segue a estrutura proposta pela OGC.

Para que se possa entender como funciona a arquitetura do sistema, apresentado, é necessária uma descrição dos componentes que o formam.

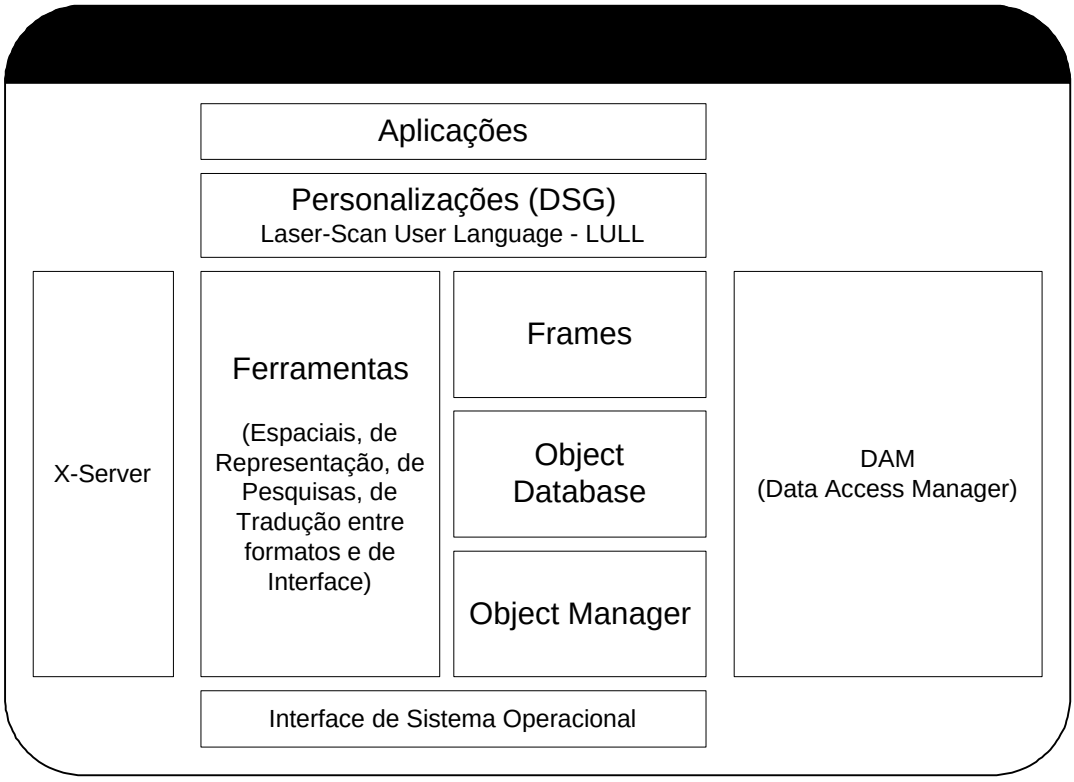


Figura 2.34 - Arquitetura do sistema *Gothic*.
Fonte: DSG, 2002.

Interface do Sistema Operacional

A Interface do Sistema Operacional está na base da estrutura *Gothic*. Se qualquer parte das ferramentas *Gothic*, *Frames*, *Object Databa*se ou *Object Manager* precisarem acessar o Sistema Operacional, então ela chama um de seus módulos internos. Se o *Gothic* é usado sobre uma plataforma diferente é necessária uma Interface do Sistema

Operacional compatível, porém o resto do sistema permanece o mesmo, independente da plataforma.

A DSG utiliza uma versão do *Gothic* cuja Interface do Sistema Operacional é baseada sobre o *Excced*, um emulador do Sistema Operacional *Unix*, que por sua vez roda sobre Sistemas Operacionais baseados em tecnologia *Windows*.

X-Server

O *Gothic* usa o sistema de gerência de janelas OSF/Motif no topo do sistema *X-Window* para todas as saídas gráficas e interações com os usuários. Isto significa que todas as aplicações baseadas em *Gothic* têm aparência e interface com os usuários consistentes. No Windows NT, também é usado o sistema *X-Window*, dando ao *Gothic* a aparência de janelas.

Há no sistema ferramentas de interface com os usuários que provêem interfaces básicas para tarefas simples, como entradas de dados, caixas de diálogos, seleções em listas, etc. Estas ferramentas comuns garantem interfaces com o usuário consistentes entre as aplicações *Gothic*.

Data Access Manager

A DAM (*Data Access Manager* – Gerenciador de Acessos de Dados) é um programa isolado que gerencia os conjuntos de dados de um ou múltiplos BD.

Sua função é validar o acesso aos BD tanto quantitativamente quanto qualitativamente, em relação ao número de clientes concorrentes e aos tipos de softwares e usuários que acessam, respectivamente.

Interface de Programação de Aplicação

Este nível da arquitetura *Gothic* provê aceso a todas as ferramentas do sistema. Por meio do LULL (*Laser-Scan User Language* – Linguagem de programação para Usuário Laser Scan) pode-se executar praticamente todas as ações do sistema, inclusive criando-se novas aplicações ou personalizar as já existentes, como o *Lamps2*. Este último processo foi o adotado pela 1ªDL quando do primeiro contato com a *Laser-Scan* e vem sido mantido até a atualidade.

Aplicações

O nível mais alto da arquitetura *Gothic* é o das aplicações dos usuários, as quais são escritas em LULL. Este é o nível fundamental do ambiente, se o conjunto das ferramentas *Gothic* são o coração do ambiente este é o cérebro, pois é neste espaço da estrutura que a personalização acontece. São os códigos escritos em LULL que automatizam os processos, definem a aplicação que usa e processa os dados geográficos.

No entanto a estratégia de desenvolvimento adotada pela DSG foi a de personalizar o *Lamps2*, aplicativo híbrido com funcionalidades de um CAD e um SIG, para processar em um só programa todas as tarefas de pós-validação e edição de Cartas Topográficas.

Banco de Dados

O Banco de Dados Geográfico Orientado a Objetos do ambiente *Gothic* está armazenado numa estrutura de diretórios dentro da pasta *data*, no servidor. Organizados em bancos e conjuntos de dados as informações estão estruturadas em formato proprietário dentro de arquivos binários.

Personalização

A estruturação dos códigos adicionados ao Lamps2 está organizada em frames, que por sua vez são armazenados numa estrutura de diretórios, com as seguintes regras:

- todos os códigos relativos aos *frames* serão armazenados numa estrutura de diretórios internos ao diretório *custom* da pasta *data*, separada em duas pastas, *lull* e *icode*, para arquivos não-compilados e compilados respectivamente;

- todo o código de um frame em particular é armazenado numa pasta com mesmo nome seguido do sufixo frame, por exemplo, o antigo *frame* 1ªDL usado nos início dos trabalhos de implementação passou a chamar-se DSG, e todos os seus códigos não-compilados ficam armazenados na subpasta *dsgframe*;

- o arquivo *_custom_frame.lull* é armazenado na raiz da estrutura, como referência para todos os outros: *dsgframe, dsgpltoframe, dsgplotprocframe, dsgprocframe, dsgwebframe, jogosframe, mapplotframe* e *timerframe*.

2.4.3.2 Conceitos fundamentais do sistema

DataBase* e *Datasets

O dado usado por qualquer aplicação *Gothic* é armazenado em um conjunto de dados (*Datasets*). No interior de cada banco de dados (*DataBase*) existe um grande número de *datasets.*

Um banco de dados numa aplicação *Gothic* é uma coleção de informações. Ele contém dados geográficos, tais como dados vetoriais, dados matriciais (*rasters)* ou imagens, e dados não geográficos, assim como detalhes de sistemas de projeção, conjuntos de parâmetros ou metadados.

Até 99 bancos de dados diferentes podem ser criados dentro do *Gothic*. Cada banco de dados é completamente independente de outro banco de dados, e pode ser usado para distinguir dados de diferentes projetos. Isto é útil para separar dois bancos de dados distintos, por exemplo, um para dados de todo o Brasil e outro para dados das Áreas de Suprimento Cartográfico (ASC) das Organizações Militares Diretamente Subordinadas (OMDS) da DSG.

Conjunto de Dados (*Datasets*)

Todos os dados são armazenados em conjuntos de dados, portanto cada banco de dados é constituído por tais conjuntos. Tipos diferentes de *datasets* podem ser categorizados formalmente naqueles que contêm o mundo real e aqueles que contém informação de apoio. Em termos de aplicações não há tratamento distinto entre um e outro, entretanto são diferenciados para fins didáticos.

Os datasets do mundo real armazenam dados que podem ser considerados como informações geográficas, tais como construções, rodovias, rios, elevações ou no caso de imagens de satélite a refletância de uma banda. Diferentes formatos de CD são apropriados para tipos de informações diferentes.

Os *datasets* de apoio contêm uma variedade de informações que são requeridas pelo aplicativo para mostrar na tela ou manipular os conjuntos de dados do mundo real. Por exemplo, são armazenados dados matemáticos referentes às projeções cartográficas ou

simbolizações pré-definidas pelo usuário para o tipo de representação cartográfica dos dados (edição). Esses são usados pelos aplicativos em conjunto com os *datasets* do mundo real.

Um *dataset* que compreende o mundo real ou os dados de apoio estará baseado em tipo único de *dataset* que não equivale a qualquer um mencionado anteriormente o qual é conhecido como conjunto de dados Base (*dataset* Base). Todos os conjuntos de dados do mundo real ou de apoio derivam de um *dataset* Base**.** Este contém definições fundamentais dos níveis do sistema. O *dataset* Base é criado quando o banco de dados é criado e nunca deve ser apagado. Uma descrição detalhada dos diferentes tipos de *dataset* é dada logo a seguir.

2.4.3.2.1 Tipos de Conjuntos de Dados (*dataset*)

***Dataset* do Mundo Real**

Geográfico - O *dataset* geográfico contém dados do mundo real armazenados como objetos (Por exemplo rodovias e construções), e usualmente são constituídos de um componente geométrico (um ponto, linha ou área), e atributos qualitativos ou de informação (endereço, propriedade, material que e feito etc...).

Os *dataset* geográficos podem ser considerados como tipos de dados vetoriais dentro do modelo no contexto tradicional de GIS. O termo vetor é usado em algumas *interfaces* do *Gothic* e pode ser considerado como um sinônimo do termo Geográfico.

Matricial (*Raster*) - O *dataset* matricial armazena informação em uma matriz regular de pixels que consistem em valores de número digitais. Os conjuntos de dados matriciais no *Gothic* podem ser binários (como mapas digitalizados matricialmente), código (mapas temáticos) ou tons de cinza (imagens de satélite) e podem ser de banda simples ou múltipla. Normalmente os conjuntos de dados matriciais estão relacionados a um sistema de coordenada de forma que eles possam ser exibidos juntamente com outro CD.

MDT (Modelo Digital de Terreno - *Digital Terrain Model*) - O *dataset* MDT é matricial e representa as elevações de terreno em tons de cinza no qual o valor da célula representa a altura de uma dada localização. Eles diferem dos conjuntos de dados matriciais em tons de cinza pelo fato daqueles representarem imagens bidimensionais.

Gazetteer - O *dataset Gazetteer* contém nomes de lugares com coordenadas de X e Y. Estes conjuntos de dados são usados como um livro de endereços para acessar diferentes localizações através do nome ao invés de coordenadas.

Embora sejam definidos ao nível do *Gothic*, o uso dos tipos de *datasets* diferentes é influenciado pela natureza da aplicação. Por exemplo, um *Gazetteer* pode ser usado em uma aplicação para localização de estabelecimentos comerciais (análise de estratégias de implantação de um novo empreendimento), enquanto um MDT seria particularmente adequado a uma aplicação de modelagem para difusão de sinais de ondas de rádio (análise de melhor posicionamento de torres de transmissão).

***Dataset* de Apoio**

Espaço - O *dataset* espaço contém os detalhes do sistema de coordenadas usado para manipular os dados. Nele se encontram definições das projeções cartográficas, *datuns*, transformações e ainda pontos de controle (CD de pontos de registro).

Usuário - O *dataset* usuário **a**rmazena informações sobre o ambiente de trabalho. Por exemplo, definição das representações cartográficas (parâmetros de edição), tabela de cores e definições das janelas de trabalho.

Tradutor - O *dataset* tradutor contém detalhes dos formatos de importação e exportação que são usados pelo aplicativo *Translate* no momento da transferência de dados entre um formato externo e o *Gothic*, bem como dentro do ambiente de trabalho do BD *Gothic*.

Help - O *dataset help* Contém os textos de ajuda.

Bloco de rascunho (*Scratch pad*) - Os conjuntos de dados de bloco de rascunho armazenam dados temporariamente durante a edição de uma sessão editoração.

Vários outros CD podem ser definidos em função das aplicações que estão sendo usadas.

2.4.3.3. Banco de Dados Orientado a Objetos *Gothic*

Modelo de Dados Objeto

O *Lamps2* usa o modelo de dados de objeto para armazenar e administrar dados.

Os objetos podem corresponder aos objetos de mundo real ou aos conceitos mais abstratos, tais como aqueles específicos do desenho cartográfico. Esses objetos podem ser modelados em termos das suas formas, localizações, relação com outros objetos, e outras propriedades. Os objetos podem também exibir comportamentos, definidos em termos de funções (métodos).

A vantagem do modelo objeto está em direta correspondência entre o modo no qual entradas reais são percebidas e o modelo de dados desenhado para representá-los. O DODS *Gothic Lamps2* fornece funcionalidade para gerenciamento de objetos e versões, e controla as ações de outros módulos do *Lamps2* envolvidos na criação, eliminação e manipulação de dados de objeto. Isto assegura que a integridade de um objeto seja mantida, e fornece a conexão entre o identificador do objeto e a localização de armazenagem dos dados associada com o objeto. Esta conexão é invisível ao usuário.

Esquema

O modelo de dados está contido no esquema (*schema*) do *dataset*. O *schema* é versionado do mesmo modo que os dados são versionados para manusear a mudança em um modelo de dados através do tempo. As entradas no schema definem as classes objeto, seus atributos, relações, comportamentos e definições de armazenamento.

A forma de definição de um *schema* interativo é fornecida para permitir operações fáceis de definição de esquema, tais como adicionar novas classes ou atributos de objetos, ou definição de herança de classe. As rotinas de definição de *schema* estão disponíveis no API Lull *Gothic*, permitindo carregamento de volume de *schema* a partir de definições externas. As rotinas estão também disponíveis para a modificação de *schema*, tais como eliminação de classes e valores.

Classes

Ao invés de tratar cada objeto individualmente, o DODS lamps2 assume que há alguns objetos mundo real que são suficientemente similares a cada outro que podem ser considerados para pertencer à mesma classe.

Cada *schema* para a classe determina quais valores podem ser escritos para um objeto na classe. Por exemplo, o objeto classe *building* (construção) pode ser definido para ter localização, endereço, área de pavimento, número de pavimentos/valores. Cada objeto, de uma classe de objeto particular tem o mesmo conjunto de valores como cada

outro objeto na classe.

Objetos

Em cada *dataset*, objetos são usados para representar objetos do mundo real. Objetos podem ter valor, referências, métodos e comportamentos. Cada objeto possui um único *object_id* (equivalente a chave primária em outros ambientes computacionais), mantendo-se constante, mesmo quando o objeto é manipulado ou editado.

Valores

Além do *object-id*, o objeto pode ter outros atributos, que podem ser simples tipos de dados, como números reais, inteiros, *strings* (texto) e *booleanos* (lógicos) ou pode ser tipo de dado bem mais especializado, como geometrias, imagens *raster* (matriciais), tabelas e referências para outros objetos.

Referências

Um valor que aponta para outro objeto é uma referência estrutural. No DODS *Lamps2*, todas as referências são bi-direcionais. Por exemplo: se a classe objeto construções tem um valor referência, localizando os pontos da entrada principal da construção ao longo do objeto estrada, a classe objeto estrada tem um valor de referência equivalente à entrada que aponta para uma coleção de todos os objetos construções na estrada.

Métodos e Comportamentos

Além disso aos valores das classes de objetos podem ser designados métodos que definem funções de comportamento que são executadas quando um método particular, ou mensagem, for enviado a um objeto na classe.

Composição de Objetos

O modelo de dados suporta classes de objetos que são compostas por outros objetos completamente genéricos. Esses grupos de objetos usam referências e são controlados pelo mecanismo *bundle*. São usados "Suporte objetos" pelos *bundles* para prover a coordenação espacial, ou seja, as classes filhas, componentes da agregação, devem herdar do sistema "Suporte to *bundle*".

Herança

Se o usuário quiser definir uma classe objeto nova, todos os atributos, referências, métodos e comportamentos têm que ser definidos no esquema.

Entretanto, se uma classe objeto for igual a uma classe objeto existente, é

possível criar uma nova classe que herde os valores, referências, métodos e comportamentos de outra classe. As modificações, como valores adicionais ou troca de comportamentos para métodos de herança podem ser definidos.

A classe na qual a nova classe está baseada é chamada de super-classe, e a classe nova criada que herda itens a partir da super-classe é chamada de sub-classe. Como as classes posteriores são criadas usando a herança, uma árvore de classes relacionadas, ou a hierarquia da herança será desenvolvida.

É possível para uma classe objeto ser definida para herdar valores, referências, comportamentos e métodos a partir de mais de uma classe parente. Esta é chamada de múltipla herança. Quando uma herança múltipla for usada, é possível que um valor, método ou comportamento que será herdado a partir de uma super-classe confronte com um valor a ser herdado de outra super-classe. Quando uma classe nova for definida, o usuário indica uma lista de super-classes das quais a nova classe deveria herdar valores, métodos e comportamentos. A ordem das classes na lista indica qual deveria ser herdada se um conflito acontecer.

Os valores podem se definidos para ser em foco global ou local. Um valor que foi definido para ter um foco global pode ser imputado a qualquer classe de objeto no *dataset*. Um valor que foi definido para ter um foco local pode somente ser imputado á uma classe de objeto para que esteja definida e qualquer classe que herde valores desta classe. Se um valor for definido com foco local em uma classe, é possível definir outro valor de mesmo nome, porém não necessariamente de mesmo tipo de dados, com foco local em outra na catarata.

O *Lamps2* é suprido com uma seleção de base de classes genérica a partir dos quais outras classes podem herdar. Um caso típico de classe pode ser uma linha simples, desenhada para representação de objetos de lineares do mundo real. Um projetista/desenvolvedor de aplicações, poderia criar sub-classes ´estrada´ e ´rio´. Essas classes podem ter futuras sub-classes, por exemplo, ´auto-estrada´, ´estrada A´, etc.

Ao usar a herança, acelera-se a criação de novas classes de objetos, ajudando a assegurar a indexação espacial e a definição apropriada da estrutura topológica.

Temas

O mecanismo de herança é também usado para fornecer o conceito de temas.

Esses são grupamentos de classes objetos, tais como: ´hidrografia´, ´relevo´, ´transporte´, etc. Várias formas dentro do Lamps2 que permitem o usuário selecionar uma classe de objeto, também permite um tema ser selecionado como um modo de seleção simultânea de um conjunto de classes de objeto.

2.4.3.4 Topologia no BDOO do *Gothic*

Além de empregar tecnologia OO, o banco de dados *Gothic* pode apoiar a estruturação topológica. Graus diferentes de estruturação estão disponíveis no *Gothic*, de objetos desestruturados (gráfico) até objetos completamente estruturados (simples). No caso posterior, a conectividade é armazenada conjuntamente com a estrutura primitiva, a qual é unida aos objetos do mundo real através de referências.

Necessidade do uso da topologia

A topologia e o meio pelo qual os dados geográficos podem ser estruturados a fim de possibilitar que os mesmos sejam passiveis de ser interpretados. Sem a topologia os componentes espaciais de um objeto são simplesmente gráficos, fornecendo assim apenas a geometria que representa a forma daquele objeto e nada mais. Se a interação espacial entre dois objetos e importante então a topologia e necessária. No *Gothic* as principais razões para o uso da topologia são:

- Obter informação sobre a conectividade entre objetos;
- Capturar objetos dentro de certa tolerância; e,
- Formar polígonos definidos por limites de outros objetos adjacentes.

Conectividade

Informações de como os objetos são conectados são essenciais para certas operações como, por exemplo, analise de rotas. Entretanto a informação de conectividade pode também ser usadas em situações mais simples como, por exemplo, identificar uma cidade em uma rodovia em especial. A Figura 2.35 ilustra as situações exemplificadas.

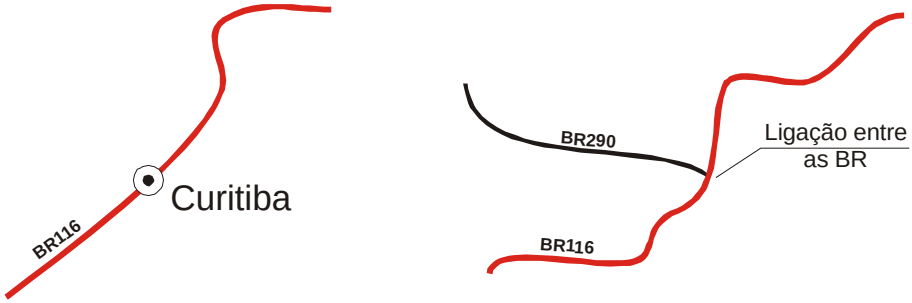


Figura 2.35 - Conectividade entre elementos lineares.

Captura

Em certas situações é importante que um objeto capture outro dentro de uma certa tolerância como é o caso dos processos de aquisição de dados vetoriais nos quais é necessário capturar um nó extremo de uma linha para dar continuidade ao desenho de uma feição. A captura e essencial para determinar a conectividade, mas é também um meio útil para manter a limpeza dos dados durante a aquisição. A Figura 2.36 ilustra um tipo clássico de situação onde a captura é fundamental.

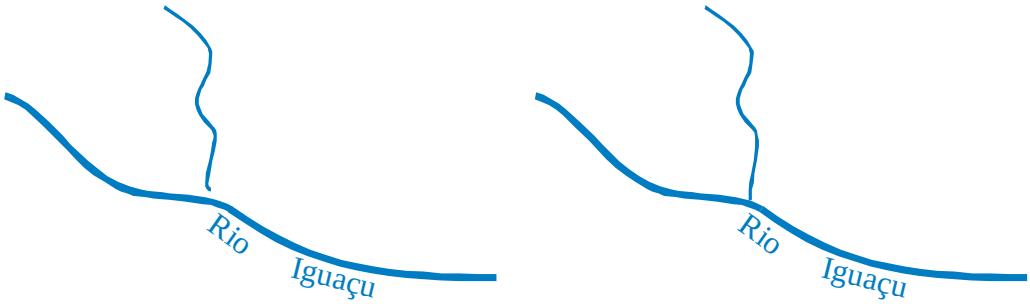


Figura 2.36 - Captura de curso d´água visando manter a integridade do sistema de rede formado.

Formação de polígonos

A formação de polígonos é o processo de obtenção de objetos tipo área por meio da união dos limites de objetos adjacentes. Para a obtenção dos polígonos é necessária que esteja garantida a conectividade. A Figura 2.37 ilustra este conceito por meio de feições comuns encontradas em cartas topográficas.

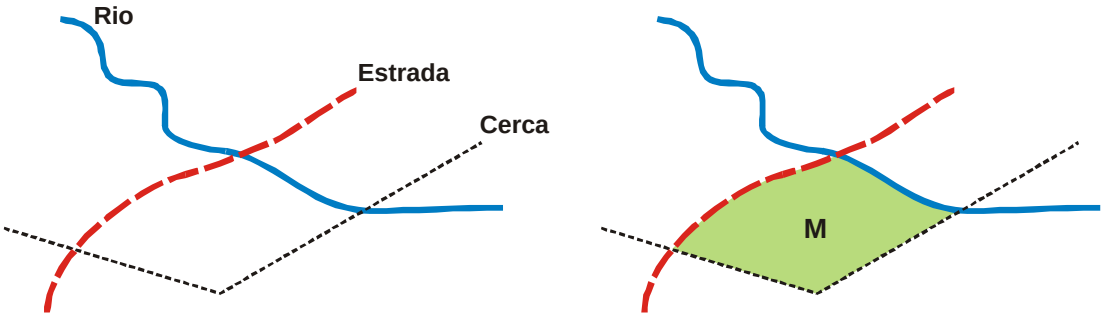


Figura 2.37 - Formação de feição de área definida a partir de feições lineares.

Tipo de Estruturas de conjuntos de dados

Nem sempre e necessária uma estrutura topológica, porém no momento da criação do conjunto de dados deve se definir se desejamos uma estrutura topológica (tipo *link-node*) ou não (tipo *spaghetti*).

Tipo *Spaghetti*

O tipo *espagheti* não suporta estruturas topológicas e pode ser empregado quando os dados já tenham sido capturados na forma estruturada com polígonos e os mesmos não vão requerer nenhuma funcionalidade de uma estrutura topológica. E o caso de um arquivo que não sofrerá atualizações. Os arquivos tipo *espagheti* são muito pouco utilizados nas aplicações de SIG.

Tipo *Link-Node*

Os dados neste tipo de estrutura suportam topologia, porém não estão obrigatoriamente estruturados de forma topológica. Dentro do *Gothic* não é possível mudar da forma *link-node* para *espagheti* e vice-versa sem a criação de um novo CD e importação.

Classes topológicas básicas

Em um CD com estrutura tipo *link-node* somente um subgrupo de classes pode requerer funcionalidades topológicas. Entretanto uma classe de objetos pode suportar topologia utilizando a herança de uma destas classes base. Existem dois tipos de classes base disponíveis, a saber, *graphic* e *simple.* Uma classe de objetos deve herdar informações espaciais de uma destas classes para armazena-las.

Graphic

Objetos que herdam desta classe base não suportam estruturas topológicas. Os dados espaciais podem ser considerados como gráficos e somente as coordenadas que definem a geometria dos objetos são armazenadas.

Ha quatro tipos de *Graphic*

- *Graphic* serve para qualquer situação, porém quando é utilizado o computador perde desempenho, é usado para as situações não enquadráveis em nenhuma das próximas ou na qual não é possível definir os limites nitidamente, por exemplo, uma zona de exclusão aérea;
- *Graphic-point*, usado, por exemplo, para ponto cotado;
- *Graphic-line*, usado, por exemplo, para contornos;
- *Graphic-area*, usado, por exemplo, para áreas urbanas.

Simple

Objetos que herdam desta classe base suportam estruturas topológicas. Informações sobre a conectividade relacionada com objetos primitivos de *links* e nós são armazenadas. Estruturando objetos *simple* e possível retomar a forma anterior livremente.

Ha quatro tipos de Simple

- *Simple* serve para qualquer situação, porém quando usado causa a perda de desempenho do computador, é usado para as situações não enquadráveis em nenhuma das próximas, por exemplo, uma construção;
- *Simple-point*, usado, por exemplo, para cidades;
- *Simple-line*, usado, por exemplo, para rodovias;
- *Simple-area*, usado, por exemplo, para áreas postais.

***Links* e Nós**

Quando uma classe base é usada a geometria é armazenada diretamente como um atributo do objeto. Entretanto quando usamos objetos *Simple* dados geográficos dos *links* e dos nós são armazenados separadamente. Para usar a topologia efetivamente, é necessário entender a estrutura *link-nó.*

Armazenamento *link-nó*

Considere o exemplo a seguir no qual duas rodovias se unem em uma junção e duas cidades ficam em uma das estradas (Figura 2.38). Quando a topologia é usada os dados espaciais são armazenados como objetos primitivos os quais são reconstruídos para a forma geométrica do mundo real do objeto quando solicitado.

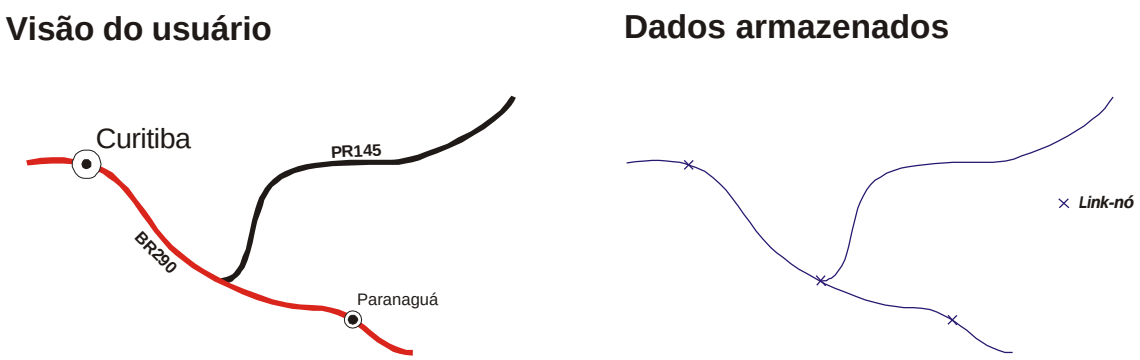


Figura 2.38 - Visão do usuário e dados armazenados no banco de dados.

Conexão entre link e nó

Uma das principais razões para usar a topologia é para definir a conectividade entre os objetos.

Esta informação é armazenada em nível de *goth-links* e *goth-nós*, essas referências são usadas para indicar como as primitivas são conectadas. Novamente, estes são usados na reconstrução ao nível de objeto para determinar como os objetos são unidos um ao outro (Figura 2.39).

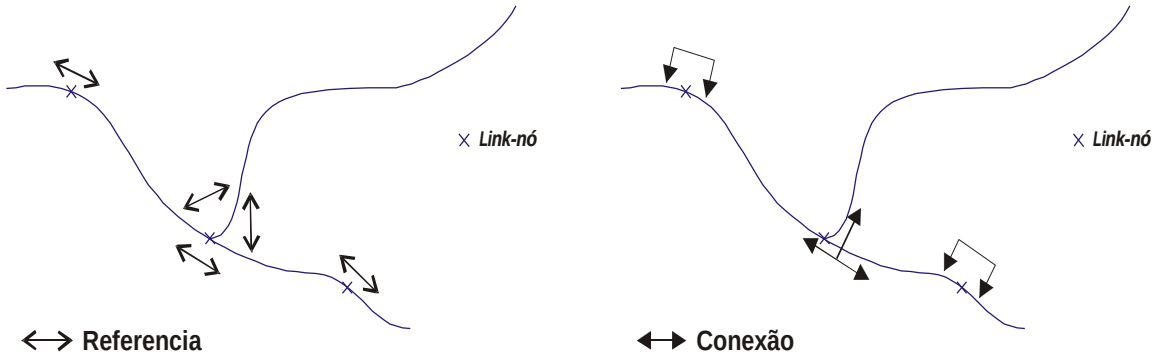


Figura 2.39 - Conexão entre link e nó para definir a conectividade entre os objetos.

Topologia *on the fly*

A topologia é mantida todo o tempo sem a necessidade efetuar processamento. Isto é possível devido a referencias usadas entre as primitivas e objetos.

Regras de topologia

Há três regras que podem ser definidas a partir de uma classe sobre outra:

- nó-compartilhado (*share-node);*
- nó-dividido (*node-split*); e
- ligação-dividida (*link-split-link*).

Estas regras serão utilizadas quando um objeto capturar outro. Sendo estas regras válidas somente para o caso de classes que herdaram características da classe base tipo *simple* devendo estar desativado o tipo *espagheti* na interface da topologia. Conseqüentemente é possível ativar e desativar essas regras. Além disso, essas regras podem ser adicionadas, apagadas ou modificadas a qualquer tempo. As regras podem ser definidas entre diferentes classes ou auto-referenciar-se para determinar como os objetos interagem com outros da sua própria classe. A Tabela 2.8 ilustra estas regras e a maneira como são aplicadas.

As regras mais complexas absorvem as das mais simples, por exemplo, nó-dividido absorve nó-compartilhado. Se mais de uma regra for definida para uma classe a que possuir maior tolerância será aplicada sobre as demais.

2.4.3.5 Conceitos Básicos de Versionamento do BD *Gothic*

Além de empregar tecnologia OO, o banco de dados *Gothic* também versiona. O princípio chave do versionamento é simples, ou seja, aquilo que muda no banco de dados, sejam nos objeto ou no nível de *schema* é que será realmente salvo. Tanto conceitualmente quando em nível prático seus benefícios são consideráveis.

Definição

A maioria das primeiras gerações de GIS e software de gerência de dados espaciais tendia a salvar dados geográficos em módulos discretos cada um dos quais tendo

um arquivo físico separado associado a ele. Sempre que eram feitas mudanças nos dados eles eram re-escritos no disco, sobre-escrevendo o arquivo já armazenado, ou salvando uma cópia completa com um nome novo. O mecanismo empregado pelo *Gothic* parte desta aproximação. O *Gothic* armazena cada mudança, tal como a definição de classes novas, a criação, modificação ou apagamento de objetos individuais. Quando a versão atual do conjunto de dados é vista, as mudanças são amalgamadas em tempo real, de forma que o usuário tem uma visão estável do conjunto de dados em seu estado atual.

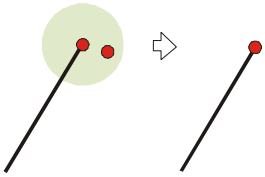
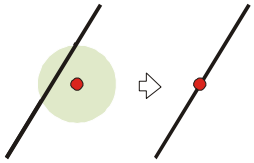
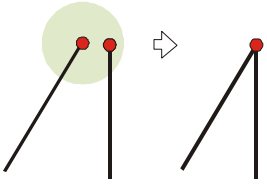
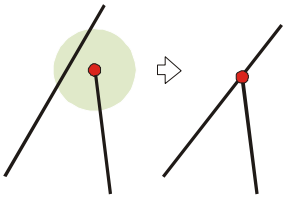
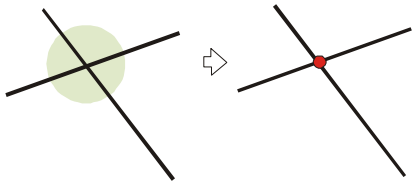
	nó-compartilhado (*share-node*)	**nó-dividido (*node-split*)**	**ligação-dividida (*link-split-link*)**
ponto x ponto		não aplicável	não aplicável
ponto x linha			não aplicável
linha x linha			
Tolerância **Nó**			

Tabela 2.8 - Regras de topologia existentes no *Gothic* (iterações possíveis entre primitivas).
Fonte: Laser-Scan, 2000.

Implicações

O fato de que somente as mudanças são armazenadas implica em inúmeros benefícios, tais como:

- possibilidade de retomar a versão anterior do conjunto de dados (*rolling back*);
- possibilidade de ter cópias múltiplas de um dataset sem aumentar a quantia de informação armazenada significativamente (*multiple copies*);

? possibilidade de dar acesso a usuários múltiplos sendo que cada um pode ter suas próprias versões para escrita (multiple user);

? transações longas podem ser executadas criando-se novos ramos das versões para estas transações (*long transacations*); e,

? os conjuntos de dados são contínuos, pois não há necessidade de subdividir em arquivos para permitir o armazenamento (*contínuous datasets*).

Cada um desses pontos será examinado detalhadamente a seguir.

Pontos de Checagem e de Retorno (*Checkpoints* e *Backtracking*)

Arvore de versões

Uma boa maneira para visualizar um *dataset* é como uma arvore de versões. Abaixo a Figura 2.40 ilustra um exemplo onde cada círculo representa uma versão.

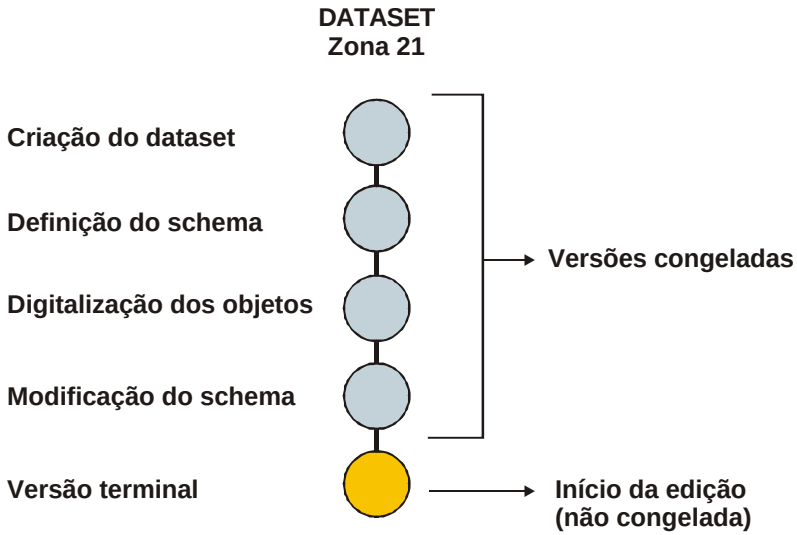


Figura 2.40 - Árvore de versões simples.
Fonte: Laser-Scan, 2000.

Quando esta sendo feita a edição do *dataset* esta edição passa a ser armazenada na versão terminal. Não é possível editar qualquer versão anterior isto previne inconsistências. Assim se dois operadores trabalharem separadamente sobre um mesmo arquivo cuja versão original designada aqui por V_n ocorrerá que quando o 1º operador salvar seu trabalho a versão a ser salva passará a ser denominada V_{n+1}, assim sendo o 2º

operador só conseguirá salvar suas alterações sobre V_{n+1} surgindo a partir deste momento V_{n+2}.

Pontos de Checagem (*Checkpoints*)

Se uma versão final e solicitada para ser editada, automaticamente é criado um ponto de checagem. Quando um ponto de checagem é estabelecido a versão corrente (final) é congelada e uma nova versão final é criada. Um ponto de checagem pode ser visto como backup incremental do *dataset*. Embora o ponto de checagem seja a ação para criar uma nova versão final, o termo é usado às vezes em referência a versão congelada do *dataset*.

Haverá situações nas quais o ponto de checagem será criado automaticamente. Por exemplo, quando o *schema* é alterado pela edição de uma nova classe ou valor. Conseqüentemente cada classe ou valor. Conseqüentemente cada versão terá sua própria versão do *schema*.

Pontos de Retorno (*Backtracking*)

Uma vantagem de criar pontos de checagem é que as mudanças na versão terminal podem ser abandonadas por *backtracking* para retomar a versão anterior, a qual será aberta como versão final. Portanto supondo-se a edição do *dataset* para adicionar um número de objetos. Se depois de efetuar as mudanças ele quiser desfazer basta efetuar um *backtracking*, que para todos os fins é como se nada tivesse ocorrido. Conceitualmente o *backtraing* é o oposto do ponto de checagem. O *backtraing* equivale ao um *undo* para a edição.

Compressing

Para acessar uma versão corrente todas as anteriores são percorridas pelo sistema, pois as versões são incrementais. Logo se o número de versões de um *dataset* é grande o acesso e as pesquisas tomam-se lentas. O *compressing* é um meio pelo qual uma série de versões pode ser amalgamada. Quando uma compressão é feita uma nova versão é criada contendo as mudanças de todas as versões escritas. Uma vez terminada as versões originais são apagadas da arvore de versões. E possível comprimir apenas uma parte da árvore de versões. Tecnicamente não há limite para o número de versões, porém é

recomendado que não ultrapassem de 10 pontos de checagem, também é recomendado que se comprima no máximo 5 pontos de checagem por vez.

Conjunto de dados Filho (*Childs*)

Denomina-se de *child* a uma cópia lógica do *dataset*, ou seja, uma cópia sem a replicação em disco do *dataset*. A Figura 2.41 ilustra um exemplo de criação de *dataset* filho para uma aplicação militar.

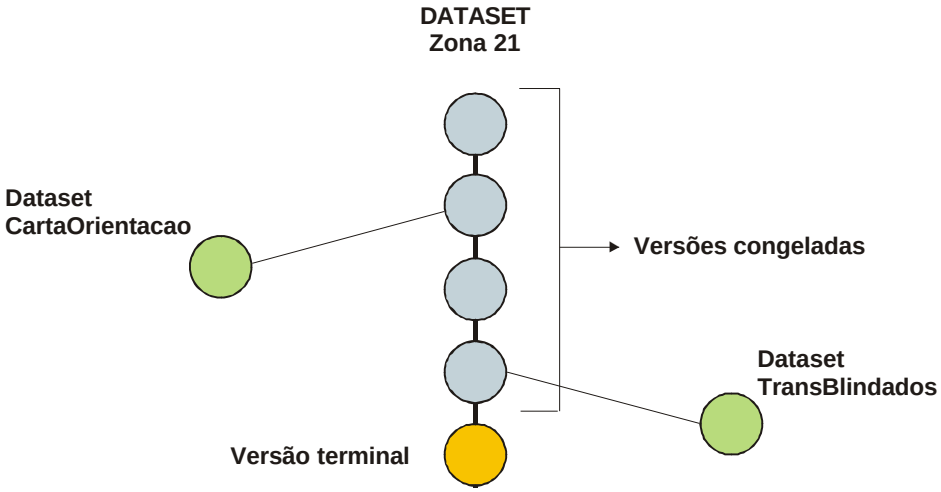


Figura 2.41 - Criação de *dataset* filho para aplicação militar.

Quando se cria um filho a versão do pai é congelada para manter a integridade entre os datasets. Se o pai for editado novamente será criada uma nova versão final deste e qualquer edição nele não será visível para o filho.

É possível criar filhos a partir de qualquer ponto de checagem de uma árvore de versões, ou seja, de versões já congeladas. Logo não ha limite para o numero de filhos ou de filhos de filhos.

Acesso a múltiplos usuários

O versionamento do BD *Gothic* permite acessos a múltiplos usuários por diferentes caminhos. Na sua forma mais simples permite dois ou mais usuários visualizar simultaneamente um mesmo BD, sendo que a um dos usuários é permitido também editar o BD.

Em uma árvore de versionamento de um BD com uma linha simples e possível a dois ou mais usuários terem acesso para leitura. Se um desses usuários iniciar uma sessão de edição então uma nova versão é criada (um ponto de checagem é estabelecido pelo sistema) para esse usuário é permitido editar e armazenar as alterações realizadas. Os demais usuários continuaram vendo a versão anterior congelada e se algum destes tentar editar o sistema prevenirá que as edições não surtiram efeito, pois a versão esta congelada.

Uma vez que o usuário que estava editando tenha concluído as operações e retornado ao modo de leitura, os usuários que já estavam vendo a versão anterior deverão executar uma atualização da visualização (*refresh*) do *dataset* para ter disponível a nova versão. Assim podemos concluir:

- dois ou mais usuários podem visualizar um BD ao mesmo tempo;
- se um usuário iniciar uma edição será criada uma versão mais nova e a antiga será congelada; e,
- outros usuários serão prevenidos da impossibilidade de edição se o tentarem.

Edição por múltiplos usuários

Em situações de produção é essencial que múltiplos usuários tenham acesso simultâneo e prolongado para edição em um BD, o termo usado para designar essa situação é Transações Longas. Para estes casos o sistema *Gothic* mantém a integridade do BD prevenindo conflitos entre diferentes usuários com a disponibilizarão de segmentos e ramos do BD.

Transações Longas

Para melhor exemplificar pode-se fazer uma comparação com as Transações Curtas em banco de dados convencionais não espaciais. A transação curta é uma edição que pode se submetida de forma imediata ao BD. Por exemplo, a retirada de dinheiro de um terminal eletrônico de banco ou e emissão do bilhete de vôo. Ambos os casos podem ser considerados para todos os fins instantâneos, e assim é possível trancar um campo do banco de dados evitando que outro editor momentaneamente execute e submeta uma mudança no BD. Impedir dois registros para o mesmo assento de linha aérea não e um problema, pois

pode ser negado a outros usuários acesso para ler e escrever àquele registro até que aquela mudança seja submetida.

As transações longas requerem um longo tempo para serem executadas e normalmente são efetuadas sobre dados espaciais. Por exemplo, a construção de uma nova rede de rodovias. A rede de rodovia somente será completada quando todas as rodovias tiverem sido construídas. Porém a duração da transação entre seu início e término pode ser considerável, e conseqüentemente a captura de dados e seu subseqüente submetimento ao BD não e instantâneo. Um mecanismo, porém é necessário para habilitar diferentes partes do BD para a edição para períodos longos, provendo ainda simultaneamente uma visão clara do estado estável do BD antes da transação começar.

Em gerações anteriores de GIS, o único mecanismo para obter a edição por usuários múltiplos era copiar um *dataset* fisicamente, e editar cada cópia. Porém, aparte das implicações práticas de armazenamento de dados, não há nenhum mecanismo parar impedir edições conflitantes em cópias distintas de um mesmo *dataset*. Desta forma esta solução está então longe de ser ideal.

O versionamento existente no sistema *Gothic* facilita as Transações Longas, pois permite para cada usuário editar um filho da versão, ou RAMO. Quando um filho é criado um SEGMENTO do pai deixa de ser disponibilizado para os demais usuários os quais só poderão obter RAMOS não coincidentes com o segmento o qual já foi criado um ramo. Enquanto isso o pai permanece como uma versão legível estável. Os conceitos de segmentos e ramos serão abordados a seguir.

Segmentos

Um segmento é definido como uma parte do *dataset*. Quando existir a necessidade de edição é possível assim definir áreas (segmento de área) para cada operador ou então disponibilizar determinada classes de objetos em um segmento. Desta forma dois operadores nunca agiram sobre os mesmos objetos criando inconsistências na integridade do sistema.

Ramos

Uma vez que tenha sido definido o segmento, a edição por usuários múltiplos é alcançada pela criação de um Ramo (*branch*) do *dataset* usado para definir o segmento. Um ramo é essencialmente um caso especial de filho de um *dataset* pai. Quando o ramo é criado o segmento não pode ser reutilizado mais, isto visa impedir que edições conflitantes ocorram.

A edição pode então ser iniciada para aquele ramo do *dataset*. O usuário de um ramo estará habilitado a ver todos os objetos do *dataset* pai, e poderá editar qualquer um deles livremente, incluindo os que estão fora do segmento definido. Entretanto ele só conseguira fundir (*merge*) novamente com o pai os que estiverem definidos no segmento. Pontos de Checagem, *backtracks* e compressões podem ser operadas no ramo como se fosse um *dataset* normal.

Merging

O processo de atualização (*update*) do *dataset* pai usando o material do ramo editado chamamos de fusão (*merging*). Quando o ramo é fundido retomamos ao pai, e somente edições definidas no segmento original serão fundidas. Qualquer das edições fora do segmento não será fundida.

Somente apos todos os ramos serem fundidos com o pai e que este poderá ser novamente editado.

Mais sobre *Merging*

Um ramo pode ser fundido apenas uma vez. Se tentarmos fundir uma segunda vez não será possível faze-lo. Portanto é recomendado que se apague os ramos após a fusão, para evitar futuras confusões. Se quisermos podemos recriar o ramo novamente a partir do segmento original.

E possível criar ramos dos ramos, porém deveremos refundi-los com o ramo original antes que este seja fundido com o pai.

Clones

Podemos usar ramos para fazer simulações de cenários usando o que chamaremos de clones. Um clone é uma copia de um ramo. Entretanto somente um clone poderá ser fundido novamente com o dataset original.

A possibilidade de investigar cenários e muito útil para se efetuar simulações procurando a melhor solução para um problema e quando a melhor solução é encontrada as demais são descartadas (eliminação dos demais clones).

Capítulo 3 - METODOLOGIA

3.1 Validação de banco de dados geográficos

A validação de banco de dados geográficos é o processo que determina se os valores do banco de dados estão razoavelmente exatos, completos e logicamente consistentes. “Razoavelmente exato” dependerá da proposta de uso dos dados, tanto valores para um estado relativo e padrões requeridos são geralmente mais apropriados que dados com exatidão absoluta.

A validação geralmente consistirá de vários passos, incluindo checagens lógicas, avaliação de exatidão e análise de erros. Neste contexto, a avaliação de exatidão é a determinação da exatidão relativa espacial e temática para um padrão conhecido (Goodchild 1995), enquanto que a análise de erros envolve a avaliação de dados com resguardo para medir o grau de incerteza (Taylor 1982) incluindo fontes de erros, erros de uso e erros de processamento (Beard 1989).

3.1.1 Propósito

Uma forma simples de validação é uma análise direta dos próprios dados; até certo ponto os usuários potenciais são capazes de determinar se os dados em questão são apropriados para uma determinada aplicação por meio da leitura de documentação que acompanha os dados (metadados). A análise dos metadados de novos dados é geralmente necessária para a validação dos dados existentes.

Isto pode ser pensado como duas fases num processo, com uma validação primária examinando diretamente propriedades verificáveis dos dados e uma validação secundaria envolvendo vários níveis de julgamento dos usuários durante a avaliação.

Em alguns casos a validação primária será desnecessária, por exemplo, quando o metadado informar o usuário que certas formas de consistência lógica foram validadas durante a construção inicial dos dados. Exemplos de validação primária e secundária são listados no Quadro 3.1.

Validação Primária

- consistência cartográfica lógica
  - polígonos fechados
  - denominação única para cada polígono
  - inexistência de arcos duplicados
  - inexistência de arcos prolongados (pontas livres)
  - inexistência de segmentos curtos
  - feições similares usando símbolos similares
- consistência lógica de atributos
  - valores dentro de uma variação lógica (exame de valores ilegais)
    - datas (e.g. meses menores ou iguais a 12)
    - hora do dia menor que 24:00 horas
    - nome de data cadastrada fora do formato adotado
    - valores de precipitação iguais ou menor que zero
  - vinculação de feições com campos de atributos
    - em um mapa de polígonos que representa lagos está incluído dados de profundidade e salinidade para cada lago?

Validação Secundária:

- pergunta lógica e testes estatísticos para os dados espaciais e atributos (procura por valores improváveis)
  - pontos colocados em posições distantes no mapa
  - elevações com valores razoáveis
- verdade de campo ou comparação a padrões conhecidos
  - coleta de amostra de áreas de campo e comparação com o banco de dados
    - avaliação da exatidão espacial
    - avaliação da exatidão de atributo
  - integralidade dos dados ("modelo tipo " de integralidade - relativa às necessidades do usuário)
    - em um mapa de estradas, todas as estradas importantes para os usuários estão incluídas?
- análise de sensibilidade
  - mude os dados, e veja se as mudanças afetarão os resultados para sua aplicação.

Quadro 3.1 - Níveis de validação primária e secundária de um banco de dados.

Os principais tipos de validação, listados no Quadro 3.1, mostram que um tipo singular de dado pode ser validado de várias formas diferentes. Por exemplo, suponha que você tenha um mapa de pontos com dados de elevação como atributo. Você pode automatizar a performance da validação num software de SIG que assegure que cada ponto é desenhado com o mesmo símbolo e que a elevação para cada ponto possua um valor de inteiro.

Estes são alguns exemplos de validação primária as quais devem ser (mas nem sempre são) executadas pelo *software* de SIG. A validação secundária do mesmo ponto poderia incluir testes estatísticos para enxergar valores discrepantes, por meio de consultas sobre os dados para determinar se o valor da elevação cai dentro de uma variação razoável ou checar para ver se todos os pontos ocorrem dentro de uma tolerância razoável.

3.1.2 Responsáveis pelo processo

O ideal é que o produtor dos dados execute uma validação básica dos dados e documente-a no metadado. Portanto, cada usuário dos dados deve conduzir algum nível de validação antes de usa-los, para assegurar a adequação dos mesmos em sua aplicação. Isto pode envolver somente uma rápida vistoria geral dos sumários estatísticos da base de dados, ou uma revisão da documentação do metadado. Em muitos casos, entretanto, esforços mais extensivos de validação podem ser necessários, requerendo uma avaliação mais profunda dos dados.

3.1.3 Necessidade de validar os dados

Atualmente é consenso geral que os dados geográficos devem obrigatoriamente ser preparados antes de serem usados nos SIG. O preparo dos dados implica em valida-los de acordo com os princípios básicos estabelecidos não só por regras topológicas, mas também por aquelas oriundas dos conceitos relacionados ao uso de BD. Alguns aspectos podem ajudar a entender os principais motivos para a necessidade de validar os dados.

a. Dado digital

Os produtos tradicionais de dados tais como mapas topográficos do governo são baseados em formatos analógicos conhecidos com padrões de exatidão publicados. Os usuários dos dados são submetidos freqüentemente a exatidão espacial e temática destes produtos, e estão cientes das necessidades cartográficas do aumento crescente da espessura de linha, da simbologia para representação do objeto e do grau de generalização com as mudanças de escalas dos mapas.

Quando nos distanciamos dos produtos analógicos para o domínio digital, a qualidade dos dados torna-se menos intuitível e menos compreendida, necessitando maior demanda por parte dos usuários para avaliá-los antes de usá-los em uma aplicação.

Os dados digitais são também mais complexos do que muitos produtos analógicos, mantendo freqüentemente ligação não tão direta com as informações temáticas subordinadas, e esta complexidade aumentam a necessidade para a melhora da compreensão dos dados. Compreende-se geralmente que todos os mapas são generalizações, mas quando séries múltiplas de dados espaciais e não espaciais são combinadas em uma camada de mapa no SIG, uma compreensão da qualidade dos dados transforma-se um elemento essencial para avaliar o valor de um banco de dados.

b. Aumento de compartilhamento e produção descentralizada de dados

Em adição ao aumento da demanda por dados espaciais, a portabilidade e a acessibilidade dos dados estão permitindo que mais pessoas publiquem seus próprios dados, quando comparado à distribuição centralizada mais tradicional de produtos analógicos. Com mais usuário criando séries de dados para aplicações específicas, está ocorrendo o aumentando do potencial de enganos que compromete a qualidade dos dados, bem como a aplicação inapropriada dos dados. Em muitos casos, os usuários podem ser guiados pelo metadado que acompanha a série de dados; entretanto em outra situação o usuário necessitará validar os dados para que se assegure que os mesmos estão sendo aplicados apropriadamente.

c. Maior custo de produção

Mesmo os usuários dos dados que geraram sua própria informação estão freqüentemente incertos da exatidão e da consistência dos dados. Com aumento da produção em casa de dados designados para aplicações específicas, os usuários se tornaram responsáveis pela sua própria validação. Os resultados da validação são relatados freqüentemente como a parte do metadado, um componente de dados requerido para muitos projetos.

Dentro deste panorama os produtores de dados geográficos devem se preparar para atender a expectativa dos usuários em relação à garantia de uma qualidade mínima para o fornecimento de dados espaciais. As instituições responsáveis pelo

mapeamento sistemático em sintonia com estas necessidades procuraram desenvolver tecnologias capazes de realizar o tratamento dos dados geográficos dentro do processo de confecção cartográfica.

3.2 Estruturação e validação de dados vetoriais para SIG

Conforme comentada anteriormente, a necessidade de arquivos para SIG tem sido o principal fator para produção deste tipo de dado. Dentro deste contexto a DSG, responsável pela produção de parte das cartas do mapeamento sistemático do território brasileiro teve de adaptar-se a nova realidade implantando nas Divisões de Levantamento uma fase para atender esta demanda.

Para que o produto cartográfico da DSG possa ser utilizado em um SIG, há necessidade da manipulação dos arquivos digitais, realizando a padronização de atributos e a exclusão de erros e excessos gerados pelos processos de aquisição de dados. Essa manipulação é chamada de validação e estruturação de arquivos digitais em formato vetorial.

A metodologia de validação e estruturação de arquivos digitais vetoriais desenvolveu-se na DSG com a necessidade de se fornecer arquivos sem erros e já estruturados, aptos para um SIG, para a Companhia Paranaense de Energia Elétrica (COPEL). Para isto foi necessária a criação da Tabela da Base Cartográfica Digital (TBCD) que padronizou a aquisição e organização dos elementos espaciais cartográficos (objetos do mundo real), no *software MicroStation*. Com os objetos organizados, deu-se início aos trabalhos de modelagem do banco de dados e desenvolvimento de algoritmos capazes de automatizar processos mecânicos da cartografia, dando origem então ao que chamamos hoje de validação e estruturação. Esta metodologia criada melhorou a qualidade dos produtos cartográficos e aumentou a capacidade de produção cartográfica da DSG. A metodologia é continuamente aprimorada, com vista ao melhoramento da qualidade dos produtos bem como a velocidade de produção cartográfica.

É importante salientar que a validação dos arquivos digitais vetoriais é realizada por meio da aplicação de rotinas baseadas em algoritmos geométricos que verificam a integridade espacial dos dados. O processo é executado sobre os vetores oriundos das fases de aquisição de dados, que basicamente são dois, a restituição e a vetorização. Outro fator importante que deve ser lembrado é que atualmente a validação

faz parte do processo de automatização cartográfica se situando entre a aquisição de dados e a edição.

3.2.1 Metodologia e Objetivo

A metodologia de Estruturação e Validação de arquivos digitais vetoriais no âmbito da DSG tem por objetivos o que segue.

- ? Padronizar os arquivos digitais, tomando como base a Tabela da Base Cartográfica Digital;
- ? Simplificar a representação gráfica de elementos espaciais, eliminando exageros, como por exemplo, excessos de vértices;
- ? Criar "nós" em cruzamentos de elementos espaciais lineares;
- ? Eliminar pontas de linhas livres oriundas da aquisição ou causadas por erros durante os processamentos e manipulações;
- ? Eliminar pontos duplicados;
- ? Remover linhas duplicadas ou muito próximas, dentro de uma mesma categoria que configuram redundância;
- ? Revisar e corrigir a topologia e toponímia oriunda da aquisição;
- ? Criar todas as áreas de todas as categorias;
- ? Ligar os elementos gráficos a registros do Banco de Dados, deixando-o estruturado em tabelas, pronto para ser povoado, configurando assim um SIG.

Ao término dos procedimentos de validação e estruturação, deverão estar garantidas as conectividades e clareza dos arquivos digitais correspondentes às categorias definidas na TBCD.

3.2.2 Critérios para criação de classes de elementos espaciais

A criação de classes de elementos espaciais para fins cartográficos segue critérios bem definidos para facilitar sua manipulação pelo usuário. A organização dos dados espaciais em categorias de dados geográficos deve ser realizada de maneira clara e com regras que obedeçam ao senso comum. A seguir serão abordados os aspectos relacionados ao assunto para um melhor entendimento:

a) Os elementos espaciais de uma mesma categoria terão por característica fundamental o seu aspecto funcional para o ser humano;

b) Quando os elementos espaciais não apresentarem aspectos funcionais relevantes, serão agregados em categorias, de acordo com a forma dos mesmos;
c) Todos os elementos espaciais, definidos no T 34-700, constarão na Base Cartográfica Digital e estarão enquadrados em uma categoria.

3.2.3 Definição das categorias

A TBCD criada pela DSG estabelece a existência de 9 (nove) categorias básicas que a princípio englobam todas as feições possíveis de ser encontradas dentro do espaço geográfico brasileiro. As categorias definidas pela TBCD tomam por base os princípios descritos logo abaixo:

Sistema de Transporte

Categoria que engloba todas as vias de acesso, obras de arte e edificações que sirvam de base ou apoio para o deslocamento humano, transporte de recursos econômicos ou estacionamento temporário, ligado a essas atividades.

Infra-estrutura

Categoria que engloba a base material ou econômica, nas áreas de indústria de base, energia, mineração, extrativismo mineral, comunicação, saúde, educação, saneamento ou irrigação, ou locais onde estejam sendo desenvolvidas atividades relevantes para o desenvolvimento humano. Incluindo empreendimentos construídos ou mantidos pela iniciativa pública ou privada, visando a atender a sociedade direta ou indiretamente. Seguem alguns exemplos de construções representáveis ou não em escala que pertencem a esta categoria: Fábrica, olaria, silo, frigorífico, faculdade, madeireira, serraria, britador, moinho, etc.

Edificações

Construções humanas não classificadas nas categorias Sistema de Transporte e Infra-estrutura, voltadas aos aspectos sociais, culturais e particulares. Eis alguns exemplos que possam gerar dúvidas sobre esta categoria: Bombeiros, BPM, granja, aviário, avícola, haras, cooperativa, etc.

Limites

Categoria que engloba todos os elementos espaciais utilizados para delimitar áreas. Estarão incluídos nesta categoria elementos virtuais como, por exemplo, limites de municípios, estados ou países.

Pontos de Referência

Categoria que engloba todos os elementos espaciais que são utilizados para materializar, de forma dinâmica ou estática, posições pontuais no terreno.

Hidrografia

Categoria que engloba o conjunto das águas correntes ou estáveis, intermitentes ou regulares de uma região, além dos elementos naturais ou artificiais, expostos ou submersos, contidos neste ambiente.

Localidades

Categoria que engloba os elementos espaciais que definirão os tipos e área de ocupação humana, classificados conforme a legislação em vigor.

Altimetria

Categoria que engloba os aspectos morfológicos do terreno.

Vegetação

Categoria que engloba as espécies vegetais naturais ou cultivadas, classificadas quanto ao seu porte ou quanto ao seu ciclo produtivo, respectivamente. Antes de especificar os tipos de vegetação é importante frisar a o conceito geral de cultura no contexto da TBCD. Segundo essa tabela a cultura se refere ao plantio de espécies vegetais de maneira organizada e racional, seguindo técnicas de produção e manutenção, visando à obtenção de benefícios ao final do ciclo produtivo.

Voltando a categoria vegetação, deve-se considerar os seguintes conceitos:

? Terreno Exposto: Região sem cobertura vegetal ou extremamente espaçada;

? Vegetação Rasteira: Vegetação nativa com altura inferior a 50cm (vegetação tipo capim);

? Vegetação Arbustiva: Vegetação nativa com altura variando de 50cm a 2m;

? Vegetação Arbórea: Vegetação nativa com altura predominantemente superior a 2m;

? Brejo ou pântano: Vegetação em terra saturada de água, doce ou salgada, geralmente coberta de vegetação característica (talos e aguapés);

? Mangue: Vegetação impenetrável e com raízes aéreas que se localiza em áreas sujeitas às marés, cujo solo é uma espécie de lama escura e mole;

? Turfeira: Vegetação que se desenvolve sobre a turfa;

? Reflorestamento: É o cultivo de espécies arbóreas com finalidades econômicas, cuja espécie será extraída no final do ciclo;

? Cultura Temporária: É toda aquela que, encerrado o ciclo produtivo, implicará em extração da espécie cultivada;

? Cultura Permanente: É toda aquela em que não há extração da planta no final do ciclo, cujo objetivo econômico é a extração de seus derivados (frutos, sementes, folhas, seiva, etc.);

? Cultura Temporária em Terreno Úmido: É toda aquela que é cultivada em terreno inundado e que, encerrado o ciclo produtivo, implicará em extração da espécie;

? Cultura Permanente em Terreno Úmido: É toda aquela que é cultivada em terreno inundado e que não há extração da planta no final do ciclo, cujo objetivo econômico é a extração de seus derivados (frutos, sementes, folhas, seiva, etc.);

? Extrativismo Vegetal: É uma área de vegetação nativa e mista onde uma das espécies vegetais é explorada.

3.2.4 Fluxo geral de dados

Os dados utilizados na validação e estruturação digital têm sua origem na aquisição (restituição ou vetorização). Os dados oriundos daquelas fontes são submetidos a diversas etapas dentro do processo de validação. Após a execução deste processo, ocorre o armazenamento dos arquivos em mídia adequada e o envio para a edição automatizada. A Figura 3.1 apresenta o fluxo de dados desde da aquisição até a edição.

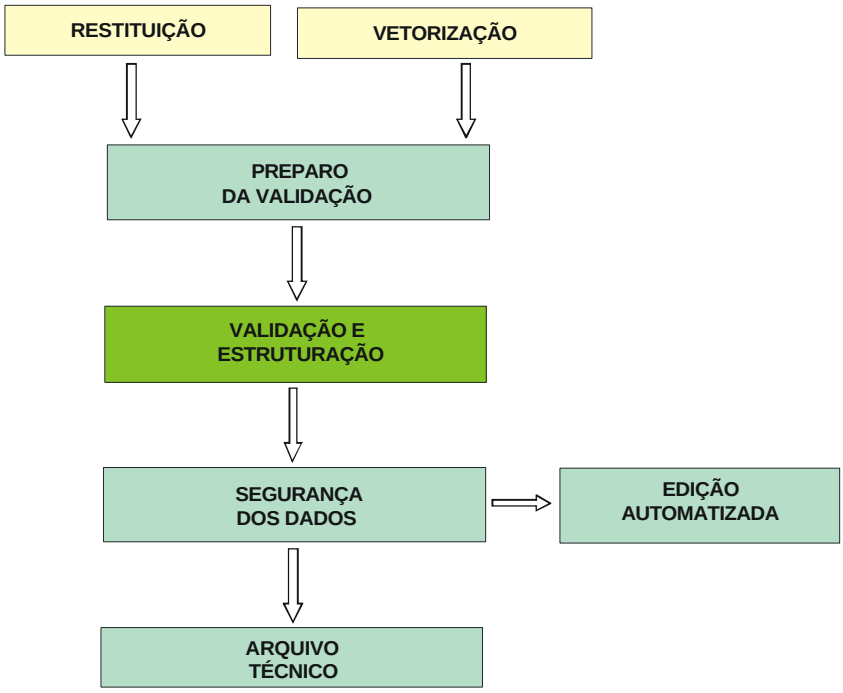


Figura 3.1 - Fluxo de geral produção de dados geográficos no âmbito da DSG, situando a validação e estruturação de dados vetoriais.
Fonte: DSG, 1ª Divisão de Levantamento.

3.2.5 Fluxo de validação e estruturação

O processo de validação e estruturação de dados vetoriais segue algumas etapas que podem ser visualizadas na Figura 3.2.

O processo de validação e estruturação de dados vetoriais utiliza diversas rotinas computacionais desenvolvidas pela DSG. O desenvolvimento destas rotinas teve início no final da década de 90 estendendo-se até início da década posterior. As rotinas foram desenvolvidas em linguagem MDL (*MicroStation Development Language*) por programadores da DSG para automatizar os processos de produção cartográfica. Uma descrição mais abrangente das rotinas citadas é dada a seguir.

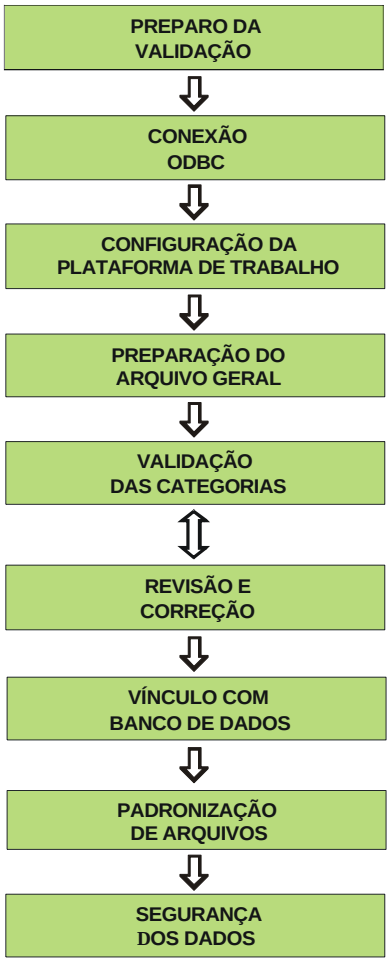


Figura 3.2 - Fluxo de validação e estruturação de arquivos vetoriais.
Fonte DSG, 1ª Divisão de Levantamento.

3.2.6 Rotinas utilizadas

Para a execução da validação e estruturação de dados geográficos no ambiente de CAD foram criados diversos programas em forma de rotinas computacionais. Estes têm por objetivo básico resolver as possíveis inconsistências encontradas nos vetores oriundos da aquisição. Os programas foram desenvolvidos em linguagens de programação distinta dando origem a dois tipos básicos de rotinas, a saber:

MDL(*MicroStation Development Language*)

São rotinas programáveis na Linguagem C, utilizadas no ambiente do *MicroStation*. A MDL é carregada de duas maneiras pelo *Prompt* do MicroStation (*Key in*) e por meio do menu *Utilities* com seguintes oprerações:

Prompt do CAD:

I. Digita-se: *MDL LOAD <nome da MDL a ser utilizada>*;

II. Pressiona-se: Enter

Por exemplo, para carregar a rotina MDL FLAGFREE basta digitar no *prompt* MDL LOAD FLAGFREE e pressionar ENTER.

Menu Utilities:

I. Seleciona-se no MicroStation: Menu Utilities > MDL Applications;

II. Seleciona-se a MDL que se deseja executar dentro da caixa *Avaliable Applications*;

III. Clicar *Load*.

A seguir é apresentado um panorama da rotina MDL CNB1 que contém um subgrupo de rotinas utilizadas no MicroStation durante a fase de validação.

MDL CNB1

A rotina foi criada para possibilitar o uso de diversas ferramentas auxiliares para percorrer e corrigir inconsistências dos arquivos de dados geográficos da aquisição. A descrição das rotinas é dada a seguir.

- **Fix Inter**: Permite a criação manual de vértices.

  Extend And Fix Endpoint: Estende e cria vértices;

  Fix Intersection: Cria vértices entre linhas que se cruzam;

- **Queue Flags**: Permite que sejam percorridos os *flags* gerados por outras MDL's

  Queue Flags Open: Carrega os *flags;*

  Queue Flags Previous: mostra o *flag* anterior;

  Queue Flags Current: mostra o *flag* corrente;

  Queue Flags Next: mostra o próximo *flag*;

  Queue Flags Get Dialog: mostra em qual *flag* se está trabalhando;

  Queue Flags Zoom First: Permite salvar o Zoom a ser utilizado;

  Queue Flags Zoom In: Permite aproximar o *flag*;

  Queue Flags Zoom Out: Permite afastar o *flag*;

  Queue Flags Close: Descarrega os *flags*;

- **Flag Elements**: Permite que sejam criados *flags* a partir de seleção prévia de quaisquer elementos.

Rotinas tipo MDL:

DDUPCELL

Identifica as células duplicadas no arquivo podendo eliminá-las automaticamente ou apenas marcá-las;

DDUPLINE

Identifica as linhas padrão *line* duplicadas no arquivo podendo eliminá-las automaticamente ou apenas alterar seus parâmetros;

DDUPTEXT

Identifica os textos duplicados ou muito próximos no arquivo podendo eliminá-los automaticamente ou apenas marcá-los;

DROPLS

Quebra todas as linhas padrão *line string* transformando-as para o padrão *line*;

EXTEND

Estende linhas a partir de suas pontas livres de acordo com os parâmetros especificados na entrada.

FIXIN

Gera interseções entre os cruzamentos de linhas no padrão *line*;

FLAGANG

Identifica e marca ângulos formados pelas linhas do arquivo seguindo uma determinada tolerância;

FLAGFREE

Identifica e marca pontas livres em linhas;

FLAGINT

Identifica e marca interseções formadas pelas linhas do arquivo seguindo uma determinada especificação;

FLAGNEAR

Identifica e marca linhas muito próximas do arquivo seguindo uma determinada tolerância;

JOINSEG

Realiza o processo inverso do programa DROPLS, unindo segmentos de reta tipo *line* dos mesmos parâmetros, onde não há interseção, transformando-as em linhas contínuas tipo *line string*;

MARKFREE

Anula as pontas livres encontradas no arquivo seguindo uma determinada tolerância;

SHORTCLE

Identifica pequenos elementos lineares segundo parâmetros especificados na entrada podendo eliminá-los automaticamente, marcá-los ou apenas alterar seus parâmetros;

SHORTSEG

Identifica e marcam pequenos segmentos de reta segundo parâmetros especificados na entrada.

MACRO

São programas criados em linguagem *Basic* utilizados no ambiente do MicroStation. Assim como as rotinas MDL, a MACRO também pode é carregada de duas maneiras pelo *Prompt* do MicroStation (*Key in*) e por meio do menu *Utilities* com seguintes operações:

Prompt do CAD:

I. Digita-se: *MACRO <nome da MACRO a ser utilizada>*;

II. Pressiona-se: Enter.

Por exemplo, para carregar a rotina MACRO CRIA_CAT basta digitar no *prompt* MACRO CRIA_CAT e pressionar ENTER.

Menu *Utilities*:

I. Seleciona-se no MicroStation: Menu *Utilities* > MACRO;

II. Seleciona-se a MACRO que se deseja executar dentro da caixa;

III. Clicar *Run*.

A seguir é dada uma descrição das rotinas desenvolvidas e onde são empregadas na fase de validação.

Rotinas tipo MACRO

AUTOREPCEL

Rotina que substitui automaticamente células do arquivo utilizando-se de uma relação constante de um arquivo de textos;

CRIACAT (macro)

Cria os 18 arquivos das categorias, sendo 9 de elementos pontuais (células), lineares e de área 9 de toponímia após terem sido separados por níveis;

FLINHAC

Rotina utilizada para realizar automaticamente o processo de vínculo dos elementos lineares com o banco de dados;

FSHAPED

Rotina utilizada para realizar automaticamente o processo de vínculo dos elementos de área com o banco de dados;

SEPARAC

Separa os elementos do arquivo em níveis diversos de acordo com as categorias: elementos lineares do nível 1 a 9 e textos do nível 11 a 19;

TXCOR

Rotina que estrutura os textos dentro de seus níveis corretos a partir de sua cor;

VALIDA63

Rotina utilizada para fechar áreas percorrendo toda seu perímetro ao mesmo tempo em que altera seus parâmetros;

VALSX

Rotina utilizada para realizar automaticamente processos do MGE;

UNDROP2

Converte *lines* para *linestrings*, eliminando os elementos (*lines*) que estão abaixo da tolerância especificada.

3.2.7 Descrição das fases da validação

Na seqüência será vista uma descrição pormenorizada das fases apresentadas no fluxo ilustrado na Figura 3.2.

3.2.7.1 Preparo para validação

O preparo funciona como uma pré-fase do processo de validação propriamente dito. Nessa fase são coletados todos os materiais para dar início à validação, tais como, original de restituição ou vetorização em formato analógico e digital, conjunto de fotos reambuladas e pasta C-101 (documento de origem da carta). Também é realizado o preparo das informações de texto ou toponímias que deve ser

inserido na carta. Sendo necessário para esta ultima atividade devido a natureza da fonte de aquisição a adoção de dois procedimentos distintos.

Nas folhas oriundas de Restituição são consultadas as fotos que cobrem a área da carta e extraídos os nomes de localidades, de rios, escolas, fazendas, edificações públicas e de outros elementos que foram classificados na fase de Reambulação.

O responsável pelo preparo deve especificar o tamanho de cada elemento da toponímia a ser colocado na folha, conforme sua escala e sua extensão, baseando-se no Catálogo de Fontes utilizado pela DSG. O lançamento dos textos é feito no original de restituição impresso.

Nas folhas vetorizadas não há necessidade de verificação de fotos e especificação de elementos da toponímia, pois todos os textos já são colocados no processo de vetorização.

3.2.7.2 Conexão ODBC

A *Microsoft*, junto com alguns outros fabricantes de banco de dados projetaram o *driver* ODBC (*Open DataBase Connectivity*) para prover um padrão, consistente e simples para conexão a bancos de dados. Com o *driver* ODBC é possível simplificar o acesso de forma que o usuário não tenha necessidade de um alto grau de conhecimento técnico para poder ter acesso a diferentes bancos de dados.

Aplicações compatíveis com ODBC: *Microsoft SQL Server, Microsoft Access, Word, Excel, Powerbuilder, Borland Delphi, Cristal Pro Report Writer, Visual Basic* e centenas de outros.

No processo de validação o *driver* ODBC faz a ligação entre a plataforma de trabalho é o *Microsoft Access*, que é o banco de dados utilizado no processo.

3.2.7.3 Configuração da plataforma de trabalho

A plataforma de trabalho usada pela DSG no processo de validação é o MGE da *Intergraph*. Essa plataforma foi adotada em meados da década de 90 por fazer na época parte da solução de automatização cartográfica estabelecida pela DSG.

A criação do projeto no MGE engloba a ligação do *driver* ODBC, chamado de *Schema* no MGE e a criação de uma estrutura particular de diretórios e subdiretórios, que irão conter os dados gráficos. Esses dados consistem de arquivos gráficos, arquivos de configuração, arquivos de sistema, entre outros.

Após a criação do projeto é realizada a importação do projeto modelo, que é um projeto previamente criado, contendo as tabelas e feições previstas na TBCD. Este passo é feito apenas uma vez para cada projeto criado.

Em seguida é verificado qual o meridiano central do fuso que a folha se enquadra, qual o *Datum* utilizado pela aquisição e copiar para o diretório "*seed*" do projeto criado, o arquivo semente referente a esses dados. A DSG preparou para este procedimento diversos arquivos sementes que contêm os *Datuns* adotados para o mapeamento sistemático e os fusos UTM que passam no território nacional (Tabela 3.1).

Datum	Arquivos semente
Astro Chuá	ACH_33, ACH_39, ACH_45, ACH_51, ACH_57, ACH_63, ACH_69, ACH_75
Córrego Alegre	CAL_33, CAL_39, CAL_45, CAL_51, CAL_57, CAL_63, CAL_69, CAL_75
SAD-69	S69_33, S69_39, S69_45, S69_51, S69_57, S69_63, S69_69, S69_75

Tabela 3.1 - Datuns e arquivos semente criados pela DSG para a plataforma de trabalho MGE.
Fonte: DSG, 1ª Divisão de Levantamento.

3.2.7.4 Preparo do arquivo geral

O arquivo geral contém todos elementos vetoriais de uma folha a ser validada e precisa ser trabalhado, para posteriormente ser feita a divisão em categorias.

O MGE precisa criar um "mapa" novo a partir do arquivo geral para que possa assumir os parâmetros estabelecidos do arquivo semente.

Um "mapa" é simplesmente um arquivo de desenho criado pelo MGE, o qual contêm parâmetros gráficos e geográficos pré-definidos no arquivo semente (*seed*).

A partir deste ponto já são realizados alguns procedimentos iniciais de validação e estruração que constituem basicamente em:

I. Copiar o arquivo original da folha (restituição ou vetorização) para dentro do mapa criado;

II. Selecionar e quebrar (*drop*) todos os elementos do arquivo (exceto células e dimensões);

III. Verificar os elementos do arquivo utilizando a *mdl levtype* e se houver curvas, retirá-las com a *mdl stroke;*

IV. Unir os elementos utilizando a *mdl joinseg;*

V. Suavizar as linhas com a utilização da ferramenta de suavização de vértices *Line Weeder* do MGE que utiliza o algoritmo Douglas-Peucker (ver Capítulo 04). Sendo necessário para este passo à adoção de tolerâncias conforme a escala da carta (por exemplo 1:25.000 = 1; 1:50.000=2).

VI. Carregar e executar a *macro separac*. Essa macro analisa os elementos conforme os atributos da TBCD, distribuindo-os em níveis de 1 (um) a 18 (dezoito), e prepara para a criação dos arquivos de categorias, conforme a Tabela 3.3;

VII. Criar os 18 arquivos de categorias por meio da execução da *macro cria_cat*;

3.2.7.5 Validação das categorias

Para a perfeita execução da validação das categorias individualmente sem que ocorra perda de dados ou qualquer outro tipo de problema relacionado à geometria dos vetores a metodologia de validação adotada pela DSG estabelece uma ordem de prioridade para execução do processo. A ordem ideal de validação das categorias é a que segue:

1º Transporte

2º Infra-estrutura

3º Edificações

4º Pontos de Referência

5º Hidrografia

6º Localidades

7º Altimetria

8º Limites

9º Vegetação

Antes de proceder à validação algumas feições devem ser levadas a determinadas categorias para permitir um perfeito ajuste espacial dos vetores em vista da manutenção da integridade topológica. Estes casos são os seguintes:

- Deverão ser copiadas para a categoria Sistema de Transporte todas as barragens do arquivo Infra-estrutura visando o ajuste de estradas que porventura coincidam com seu alinhamento.

- Deverão ser copiados de categorias já validadas, todos os elementos necessários para a criação de linhas ou fechamento de áreas nas categorias subseqüentes.

Outro fator importante que deve ser lembrado diz respeito aos elementos de área grandes, como por exemplo, no caso da categoria vegetação. Nesta categoria as

áreas grandes devem ser divididas em dois ou mais polígonos para evitar que os arquivos sejam corrompidos após a execução da rotina FSHAPED, a qual é responsável pela criação do *shape* relacionado ao polígono fechado. A divisão por sua vez deve ser feita tomando por base elementos existentes na carta, tais como estradas ou rios, devendo-se evitar neste caso a escolha aleatória de feições.

Categorias	**Níveis**	
	Arquivo geométrico	Arquivo de textos
Transportes	01	11
Infra-Estrutura	02	12
Edificações	03	13
Limites	04	14
Pontos De Referência	05	15
Hidrografia	06	16
Localidades	07	17
Altimetria	08	18
Vegetação	09	19

Tabela 3.3 - Arquivos de categorias e de textos associados.

Seqüências de operações para validação das categorias

a. Todas as categorias

A seqüência de operações que devem ser executadas para todas as categorias (exceto para a categoria vegetação) é a que segue:

I. “Quebrar” todos os elementos lineares do arquivo por meio do uso da ferramenta MDL DROPLS;

II. Eliminar as linhas duplicadas (mesmas coordenadas) (MDL DDUPLINE);

III. Gerar as interseções entre os elementos lineares (MDL FIXIN);

IV. Eliminar células duplicadas e marcar células próximas (MDL DDUPCELL);

V. Unir as linhas contíguas (MDL JOINSEG);

VI. Eliminar as linhas menores que a tolerância, com 1 ponta livre (MDL SHORTCLE);

VII. Estender todas as pontas livres, usando 10m (escala 1/25.000) e 15m (escala 1/50.000) (MDL EXTEND);

VIII. “Quebrar” todos os elementos lineares (MDL DROPLS);

IX. Gerar as interseções entre os elementos lineares (MDL FIXIN);

X. Unir as linhas contíguas (MDL JOINSEG);

XI. No arquivo de saída gerado, eliminar as linhas menores que a tolerância, com 1 ponta livre (MDL SHORTCLE)

XII. Recuar para a posição original, todas as pontas livres que não geraram interseção com outros elementos, utilizando –10m (escala 1/25.000) e -15 m (escala 1/50.000) (MDL EXTEND);

Procedimentos utilizados somente para as categorias de sistema de transporte e limites:

I. Marcar todas as pontas livres por meio da rotina MDL FLAGFREE e utilizar como referência os arquivos das outras categorias para estender as linhas. (Figura 3.3 e Figura 3.4). A rotina coloca um circulo na extremidade da linha marca de maneira a caracterizar a ponta livre;

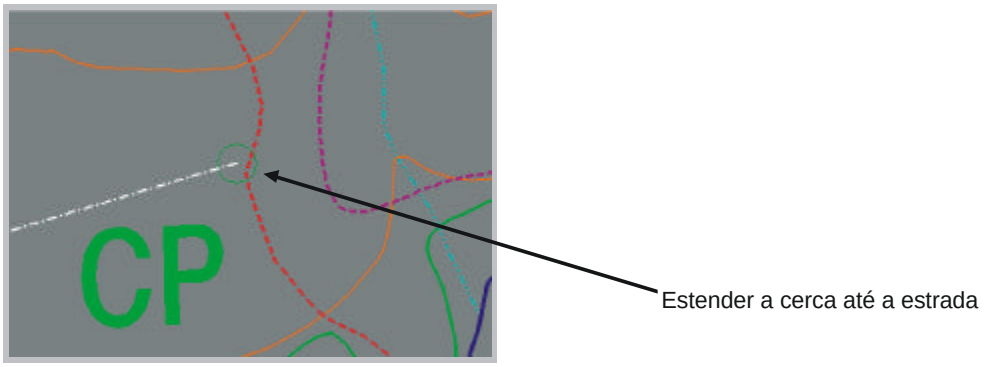


Figura 3.3 - Exemplo de feição que deve ser estendida na categoria limites.
Fonte: DSG, 1ª Divisão de Levantamento.

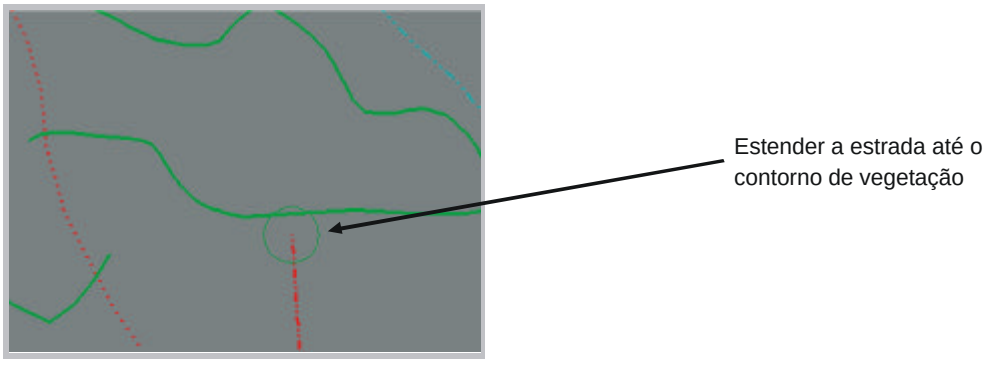


Figura 3.4 - Exemplo de feição que deve ser estendida na categoria transportes.
Fonte: DSG, 1ª Divisão de Levantamento.

II. Marcar as extremidades de linhas que devem ser livres por meio da rotina MDL MARKFREE (Figura 3.5). Esta MDL coloca um triângulo na extremidade da linha, evitando assim que a mesma seja acusada como ponta livre;

III. Percorrer e resolver as possíveis pontas livres não marcadas pelos triângulos (MDL FLAGFREE);

IV. Procurar pequenas linhas por meio da MDL SHORTCLE. Este processo visa detectar desencontros entre dois ou mais cruzamentos que poderiam se fundir em apenas um nó, conforme mostra o exemplo da figura a seguir. Com essa seleção é possível percorrer todos os elementos, eliminando somente aqueles incorretos (Figura 3.6);

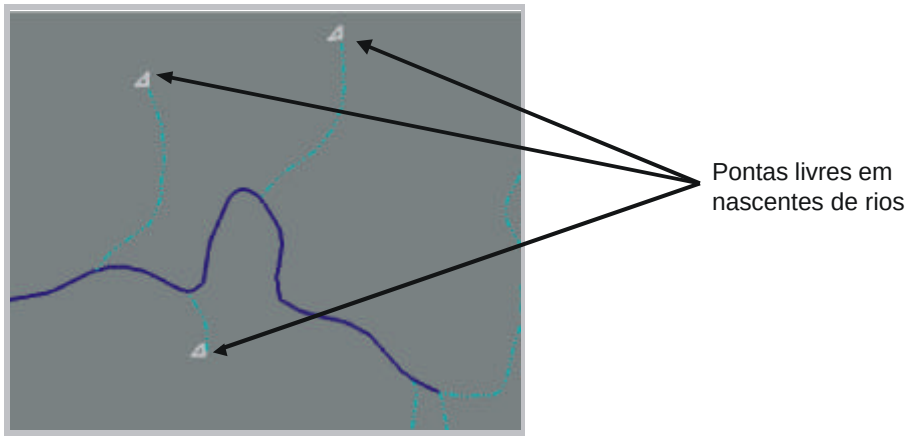


Figura 3.5 - Exemplo de falsa ponta livre na categoria hidrografia.
Fonte: DSG, 1ª Divisão de Levantamento.

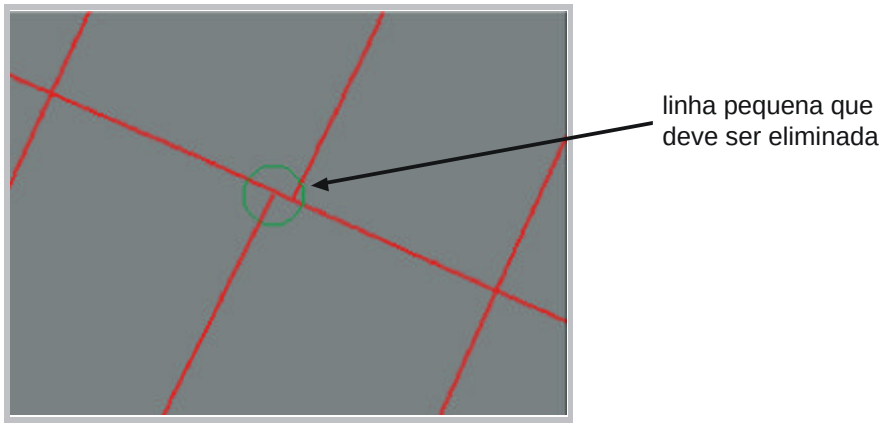


Figura 3.6 - Exemplo de linha pequena que pode acarretar erro em sistemas de rede
Fonte: DSG, 1ª Divisão de Levantamento.

V. Procurar pequenos segmentos por meio da MDL SHORTSEG. Com esta ferramenta é possível eliminar vértices desnecessários nas linhas do arquivo,

reduzindo o seu tamanho. Deve-se percorrer todas as marcas, retirando-se somente os vértices desnecessários (Figura 3.7);

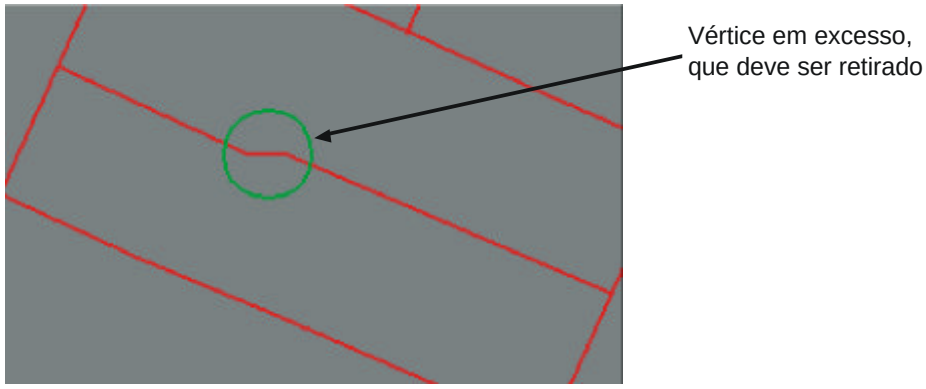


Figura 3.7 - Exemplo de pequeno segmento.
Fonte: DSG, 1ª Divisão de Levantamento.

VI. Procurar ângulos desnecessários marcando interseções entre duas linhas cujo ângulo formado por elas esteja dentro do intervalo determinado por meio do uso da MDL FLAGANG. Deve-se percorrer todas as marcas, retirando-se somente os ângulos incorretos (Figura 3.8);

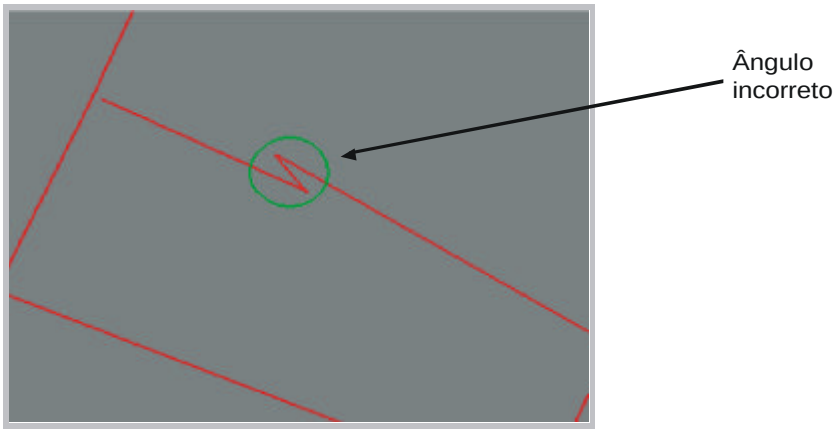


Figura 3.8 - Exemplo de ângulo incorreto.
Fonte: DSG, 1ª Divisão de Levantamento.

VII. Marcar linhas próximas, que estão abaixo da tolerância selecionada por meio da MDL FLAGNEAR. Deve-se percorrer todas as marcas, consertando-se somente os incorretos (Figura 3.9);

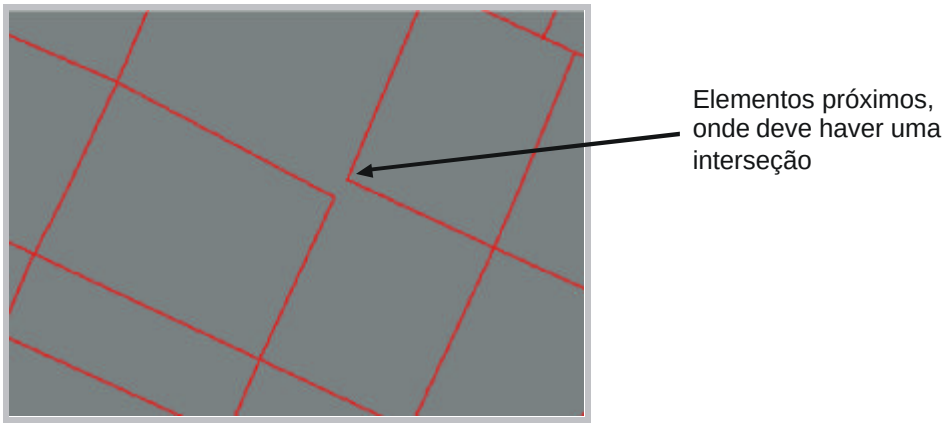


Figura 3.9 - Exemplo de elementos próximos e incorretos.
Fonte: DSG, 1ª Divisão de Levantamento.

VIII. Corrigir as interseções incorretas entre linhas (*loop*) por meio da MDL FLAGINT (Figura 3.10);

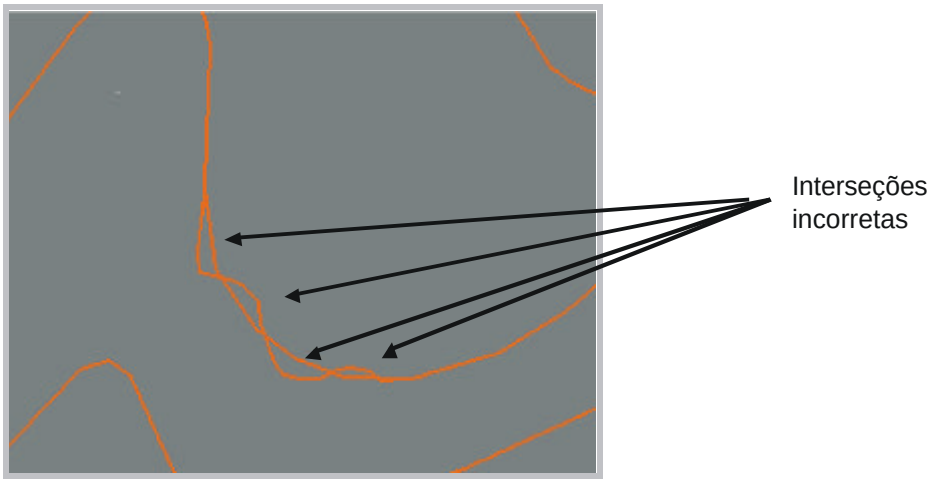


Figura 3.10 - Exemplo de interseção incorreta.
Fonte: DSG, 1ª Divisão de Levantamento.

IX. Fechar as áreas utilizando a macro VALIDA63. Esta macro possibilita que sejam percorridas as linhas que delimitam uma área, liberando um centróide identificador após o seu fechamento. As linhas processadas por meio da macro, juntamente com seu centróide, são enviadas para o nível 63 para que posteriormente seja rodada a rotina FSHAPED, que fará o vínculo das áreas com o banco de dados, transformando-as em *shapes;*

X. Marcar, percorrer e resolver as possíveis pontas livres por meio da MDL FLAGFREE;

XI. Apagar moldura, marcas de pontas livres do nível 61 e textos;

XII. Aplicar os atributos do padrão delimitador nas linhas de áreas que delimitam a folha. É importante que os cantos das folhas tenham “nós”, pois o delimitador é utilizado em procedimentos futuros de ligação de folhas;

XIII. Retirar os atributos (*links*) contidos nos elementos do arquivo por meio da aplicação da MDL TIRA_ATRIB;

Após a execução dos procedimentos para validação da categoria o arquivo gerado deve ser salvo para garantir a atualização dos dados gráficos trabalhados.

b. Categoria vegetação

Na categoria vegetação há uma grande quantidade de elementos usados para delimitar as áreas, dentre esses se incluem elementos de outras categorias. As Figuras 3.11 e 3.12 ilustram exemplos comuns de elementos pertencentes a outras categorias que geralmente funcionam como delimitadores de polígonos de vegetação.

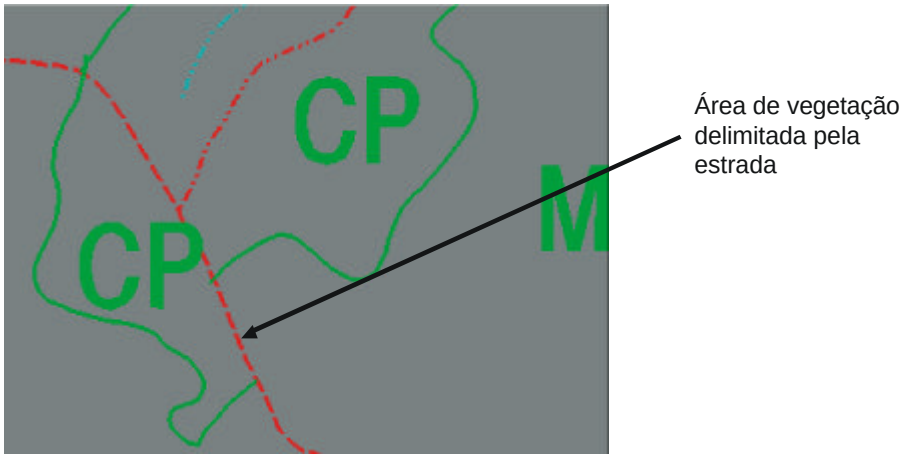


Figura 3.11 - Polígono de Vegetação limitando pela estrada.
Fonte: DSG, 1ª Divisão de Levantamento.

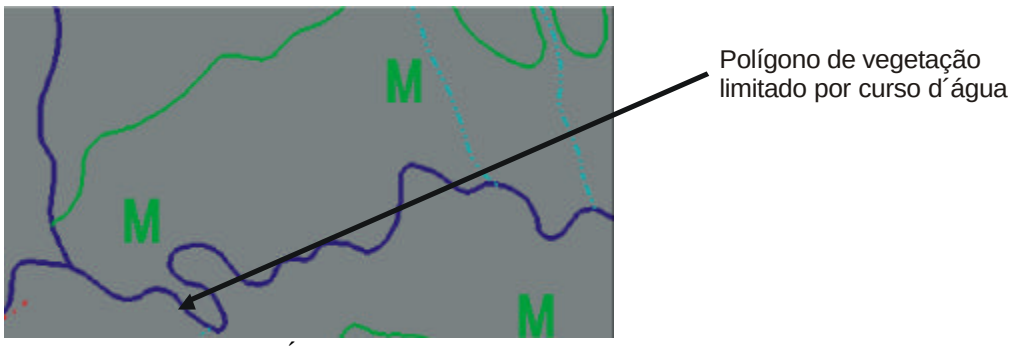


Figura 3.12 - Área de vegetação delimitada pela hidrografia.
Fonte: DSG, 1ª Divisão de Levantamento.

I. Quebrar todos os elementos lineares por meio do uso da MDL DROPLS;

II. Eliminar as linhas duplicadas usando a MDL DDUPLINE;

III. Unir as linhas contíguas por meio da MDL JOINSEG;

IV. Procurar pequenos segmentos utilizando a MDL SHORTSEG. Neste procedimento deve se estabelecer a tolerância do comprimento a serem procurados, por exemplo para escala 1:25.000 é estabelecido um comprimento mínimo de 7m e para escala 1:50.000 de 10m;

V. Estender todas as pontas livres, usando até 15m de tolerância, por meio da MDL EXTEND e apagar a moldura;

VI. Copiar das categorias Transporte, Infraestrutura, Edificações, Limites, Hidrografia, Localidades e Altimetria as feições já validadas;

VII. Apagar as feições que não serão utilizadas para formar polígonos de vegetação, por não se tratarem de delimitadores de áreas:

? Categoria **Infraestrutura**:

201 - linha telefônica;

205 - linha de distribuição de energia;

206 - linha de transmissão de energia;

219 - tubulação;

220 - tubulação;

225 - aqueduto;

? Categoria **Edificações:**

- deverão ser mantidos todos os elementos lineares e as áreas a seguir.

314 - campo de futebol;

315 - estádio;

316 - área de lazer;

317 - mercado;

318 - campo de tiro;

320 - cemitério;

? Categoria **Altimetria**:

- deverão ser mantidas apenas as áreas fechadas;

VIII. Substituir os centróides de outras categorias pelos de vegetação por meio da macro AUTOREPCEL. Esta macro tem a opção de enviar os contornos desnecessários para o nível 63 do CAD *MicroStation*;

IX. Apagar os contornos de área de outras categorias que foram enviados para o nível 61;

X. Estender todas as pontas livres em 1m por meio da MDL EXTEND;

XI. “Quebrar” todos os elementos lineares usando a MDL DROPLS;

XII. Eliminar linhas duplicadas utilizando a MDL DDUPLINE;

XIII. Gerar as interseções entre os elementos lineares por meio da MDL FIXIN;

XIV. Unir as linhas contíguas usando a MDL JOINSEG;

XV. Eliminar as linhas menores que a tolerância, com uma ponta livre, por meio do uso da MDL SHORTCLE. Nesta operação é utilizada uma tolerância de 20m para qualquer escala de trabalho para caracterizar as linhas com ponta livre;

XVI. Eliminar as linhas menores que a tolerância, com uma ponta livre, por meio do uso da MDL SHORTCLE. Nesta operação é utilizada uma tolerância de 2m para qualquer escala de trabalho para caracterizar as linhas com ponta livre.

XVII. Fechar todas as áreas usando a MDL VALIDA63;

XVIII. Apagar os elementos que não estiverem no nível 63;

XIX. Procurar ângulos desnecessários marcando interseções entre duas linhas cujo ângulo formado por elas esteja dentro do intervalo determinado por meio da MDL FLAGANG. Nesta operação é adotado um ângulo mínimo de 0º e máximo de 10º;

XX. Procurar e apagar pequenos segmentos usando a MDL SHORTSEG. A tolerância usada neste procedimento é a de 0,1m em qualquer escala para identificar os pequenos segmentos;

XXI. Marcar, percorrer e resolver as possíveis pontas livres usando a MDL FLAGFREE;

XXII. Unir as linhas contíguas por meio do uso da MDL JOINSEG;

XXIII. Aplicar o padrão delimitador nas linhas de áreas que delimitam a folha (linhas da moldura);

XXIV. Retirar os atributos (“*links*”) contidos nos elementos do arquivo por meio da MDL TIRA_ATRIB;

Seqüência de operações para os Arquivos de Texto

I. Eliminar textos duplicados e marcar textos próximos usando a MDL DDUPTEXT;

II. Passar os textos para os níveis corretos por meio da macro TXCOR;

III. Retificar parâmetros e digitar textos, com a utilização do original de restituição impresso, de acordo com a TBCD e o CATÁLOGO DE FONTES;

IV. Verificar e corrigir os atributos dos elementos dos arquivos finais usando a MDL LEVELINV;

3.2.7.6 Revisão e correção

Após a validação é feita uma impressão dos arquivos em papel para realizar a revisão da geometria dos elementos e conseqüente correção. A ligação dos elementos dos arquivos com o Banco de Dados é realizada e somente após correção dos erros detectados.

Impressão

Os arquivos das categorias são impressos em papel comum individualmente de acordo com uma tabela de cores própria para facilitar a interpretação do revisor. Os elementos de área são impressos de maneira a ter seus interiores preenchidos por cores pré-definidas para possibilitar a verificação da continuidade do espaço e adjacências de polígonos.

Revisão

As principais verificações a serem feitas durante a revisão, em todas as categorias são as seguintes:

- Verificar se todos os símbolos usados (células, linhas ou áreas) são os previstos na TBCD, para cada categoria, e se estão conforme a coluna “Representação Cartográfica TBCD”;

- Comparar com o material produzido pelo preparo, verificando se ainda existem detalhes que passaram pela fase da revisão da vetorização ou restituição e deixaram de ser representados (células, áreas, linhas e textos);

- Verificar se os elementos do tipo área estão de acordo com a legenda.

Para auxiliar no trabalho de revisão é utilizado um quadro de legendas, com as cores dos elementos do tipo área.

É interessante ressaltar que a revisão em papel das feições validadas também tem a finalidade de atestar se a geometria trabalhada não sofreu alterações e se ela se encontra com os atributos gráficos de CAD de acordo com a definição estabelecida. Problemas na continuidade do espaço representados por polígonos adjacentes neste caso podem passar desapercebidos pelo revisor da validação, pois dependendo da escala da carta trabalhada tais inconsistências podem não ser visíveis a olho nu. Portanto após esta revisão e correções, uma revisão executada é realizada. Também cabe ressaltar que em função do peso da folha a ser trabalhada (quantidade de detalhes existentes), outras impressões e novas revisões poderão ser necessárias.

Correção

A correção consiste em eliminar os erros apontados na fase de revisão. Se a folha não contiver muitos erros, a segunda revisão já poderá ser feita em tela.

3.2.7.7 Vínculo com o banco de dados

Após acontecerem as revisões em impressões e feitas as correções necessárias, os arquivos estão prontos para receberem a finalização da validação e a estruturação no banco de dados. Para realizar a ligação dos elementos gráficos com o BD são utilizadas rotinas tipo *bat*, que são operadas a partir do *prompt* do MS-DOS.

Vinculo de elementos lineares e pontuais

É feito em todas as categorias, utilizando-se a rotina tipo *bat* denominada FLINHAC. A descrição da rotina é dada a seguir.

FLINHAC

Esta rotina realiza automaticamente o processo de vincular elementos lineares e pontuais com o Banco de Dados (criação do *MS_Link*). A *interface* gráfica da rotina é ilustrada na Figura 3.13.

Vinculo de elementos de área

É feito em todas as categorias (exceto na categoria pontos de referência), com a utilização da rotina tipo *bat* denominada FSHAPED. A descrição da rotina é dada a seguir.

FSHAPED

A rotina cria as áreas usando os elementos de feições de contornos e centróides realizando automaticamente o processo de vincular estas áreas criadas com o Banco de Dados (criação do *MS_Link*). A Figura 3.14 apresenta a interface gráfica da rotina.

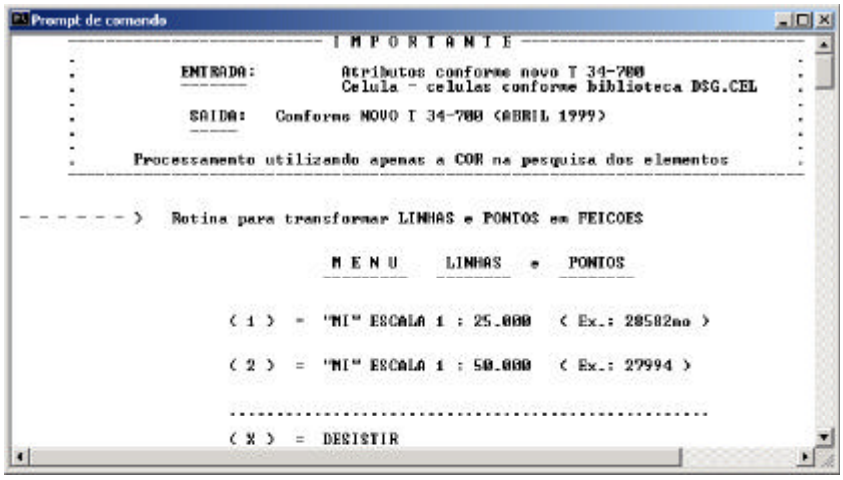

Prompt de comando
I M P O R T A N T E
ENTRADA: Atributos conforme novo T 34-700
Celula - celulas conforme biblioteca DSG.CEL
SAIDA: Conforme NOVO T 34-700 (ABRIL 1999)
Processamento utilizando apenas a COR na pesquisa dos elementos
- - - - - - > Rotina para transformar LINHAS e PONTOS em FEICOES
M E N U LINHAS e PONTOS
(1) = "MI" ESCALA 1 : 25.000 (Ex.: 28582ao)
(2) = "MI" ESCALA 1 : 50.000 (Ex.: 27994)
(X) = DESISTIR

Figura 3.13 - Janela da rotina FLINHAC.
Fonte: DSG, 1ª Divisão de Levantamento.

Prompt de comando
I M P O R T A N T E
A T U A L I Z A D A conforme T 34-700 de ABR 99
- - > Rotina para transformar AREAS em FEICOES e gerar SHAPES
M E N U SHAPES
(1) = "MI" ESCALA 1 : 25.000 (Ex.: 28582ao)
(2) = "MI" ESCALA 1 : 50.000 (Ex.: 27994)
(X) = DESISTIR

Figura 3.14 - Janela da rotina FSHAPED.
Fonte: DSG, 1ª Divisão de Levantamento.

União de arquivos de linhas, pontos e de áreas

A união dos três tipos de elementos encontrada nas cartas é realizada por uma operação simples que basicamente consiste em apagar dos arquivos de áreas criados pela rotina FSHAPED, todos os elementos que não sejam *complex shape* e copiá-los para os arquivos de linhas correspondentes.

Verificação final dos arquivos

Após a junção dos três tipos de elementos é realizada uma verificação dos elementos e correção dos possíveis erros de atributos encontrados nos arquivos finais por meio do uso da MDL LEVELINV.

Capítulo 4 – RESULTADOS

4.1 Validação em Ambiente Orientado a Objetos

Embora o termo cartografia automatizada tenha sido empregado para designar o conjunto de técnicas envolvidas na produção cartográfica assistida por computador, deve-se enfatizar que boa parte do processo de produção de dados geográficos para atender os SIG ainda requer uma grande participação de trabalho manual ou interação com o operador, quer seja na aquisição dos dados (Capítulo 2) ou na sua validação (Capítulo 3).

Para o caso da aquisição de dados atualmente diversos sistemas têm apresentado soluções de captura de informação geográfica. Como exemplo pode-se citar o caso dos sistemas que tratam da manipulação de dados de sensoriamento remoto orbital com as rotinas de reconhecimento de padrões de imagens de satélite ou ainda os sistemas de fotogrametria digital que permitem realizar a coleta de dados automáticas de informações de relevo permitindo a elaboração de modelos digitais de terreno. Algumas técnicas de digitalização vetoriais automáticas também têm se apresentado, entretanto estas ainda não se encontram em condições de uso pleno por acarretar sempre erros que sempre exigem tratamento demorado para corrigi-los.

O tratamento de dados pós-aquisição de dados vetoriais, entretanto apresenta um panorama mais favorável. Com a tecnologia existente hoje, em especial dos sistemas de banco de dados geográficos orientados a objeto (Capítulo 2), é possível tornar automáticas diversas rotinas na produção cartográfica. Isto já acontece dentro do âmbito da DSG com relação à edição automática dos dados, possibilitada por meio do uso do sistema *Gothic*.

O sistema *Gothic* permitiu indexar automaticamente a simbologia preconizada no Manual de Convenções Cartográficas T34-700 aos elementos vetoriais previamente validados em ambiente de CAD. Parte do desenho que se repetem em uma seqüência de cartas, que é o caso do mapeamento sistemático, puderam ser implementadas por programação em linguagem LULL. Em exemplo, o desenho da moldura interna das cartas, do quadriculado UTM, a malha geográfica, a disposição dos dados marginais e legenda de

acordo com a folha modelo padrão, inclusive o cálculo do ‘pé de galinha’ (relações entre o norte verdadeiro, norte de quadrícula e norte magnético) e a localização dos limites municipais na carta, sua articulação e posição em relação ao espaço geográfico nacional (estados brasileiros) foram traduzidos em linguagem LULL e implementados na rotina de edição automática.

A adoção deste sistema para realizar a fase de produção conhecida anteriormente como edição vetorial permitiu ganhos de tempo de produção extremamente significantes da ordem de 1:5. Para se ter uma idéia do ganho basta apenas comparar o tempo de produção de uma carta editada em ambiente de CAD de acordo com as técnicas convencionais que é da ordem de 150 horas de trabalho, incluindo-se aí as horas de revisão e o tempo gasto atualmente na edição automática que é da ordem de 30 horas, incluindo as revisões. As vantagens comparadas ainda em termos de versatilidade do sistema em relação ao ambiente de CAD ainda permitem que se altere sem custos adicionais de produção qualquer simbologia encontrada nas cartas armazenadas no banco de dados. Na obtenção de cartas temáticas, produto final de um SIG, esta vantagem é extremamente importante, pois através dela é possível obter maior versatilidade para visualização de fenômenos geográficos.

O potencial apresentado pela tecnologia adquirida e a ambientação no sistema possibilitaram dar continuidade aos trabalhos de automatização de processos. Dentro deste contexto procurou-se adaptar a fase de validação de dados geográficos antes executada, assim como a edição, basicamente em ambiente de CAD.

Ainda existem possibilidades de otimização de processos em fases de produção ligadas a aquisição de dados gráficos, via restituição fotogramétrica digital (LAMPS2 / SOCET SET) e digitalização vetorial de originais cartográficos (VTRAK), assuntos que deverão ser tratados futuramente em novas implementações assim que a nova tecnologia tenha sido completamente assimilada no tratamento de dados para Sistema de Informações Geográficas (SIG).

4.2 Validação de Dados no ambiente Orientado a Objetos

As possibilidades de implementações de rotinas de automatização dentro do sistema orientado a objetos permitem acrescentar fatores favoráveis a inserção de certa

inteligência aos dados geográficos. Dentro destes sistemas os dados são tratados como objeto do mundo real possuidores de comportamentos que podem ser previamente definidos numa modelagem de dados. A aplicação de regras de restrições topológica permite que de acordo com a modelagem de dados adotada os objetos geográficos tomem certas posturas em situações que afetem a sua integridade no espaço de uma carta. Esta possibilidade de modelagem de comportamentos e o uso de lógicas de programação permitiram que as rotinas criadas para o processo de validação convencional em ambiente de CAD fossem implementadas no ambiente orientado a objetos, conforme será visto a seguir.

4.2.1 Implementação das rotinas de validação

As rotinas de validação implementadas no ambiente orientado a objetos Gothic são oriundas basicamente daquelas usadas na validação de dados. São utilizados naquela fase de produção dois tipos de rotinas como mencionado anteriormente no Capítulo 4, a MDL que é escrita em Linguagem C e a MACRO escrita em Linguagem Basic. Entretanto o ambiente do *Gothic* possui sua própria linguagem de programação denominada LULL. Portanto é interessante abordar neste ponto alguns conceitos desta linguagem para ilustrar um pouco a natureza do sistema *Gothic*.

4.2.2 Linguagem LULL

LULL é a "Linguagem Usuário Laser-Scan" usada para programação no *Gothic*. Ela foi escrita pela Laser-Scan e tem sido usada para desenvolvimento de todas as aplicações baseadas na tecnologia *Gothic* Orientada a Objeto.

Desde o lançamento do *Gothic*, a linguagem LULL tem sido avaliada pelos desenvolvedores, que a tem usado para:

- escrever métodos para classes de objetos em seus bancos de dados;
- refinar a performance de datasets;
- customizar as aplicações da Laser-Scan;
- desenvolver novas aplicações baseadas sobre a tecnologia Gothic.

A linguagem LULL em si é simples, tendo alguma similaridade com o C e o Pascal. Apesar do fato da linguagem LULL não ser uma linguagem orientada a objeto

(como o Java ou o C++) ela pode ser utilizada para adicionar funcionalidade a um banco de dados. Embora isto possa ser encarado como um paradoxo, pode se fazer uma analogia óbvia com a linguagem falada, simplesmente, pois o fato de se falar uma determinada língua não impede qualquer pessoa de falar a respeito de certos assuntos.

4.2.2.1 Funções e Bibliotecas em LULL

O poder da linguagem em LULL reside nas funções. Uma função é um comando que é executado quanto chamado. Funções típicas em LULL podem imprimir uma parte de texto em um terminal (*window*) ou retornar o valor de um atributo armazenado sobre um objeto.

Funções similares são organizadas em bibliotecas da mesma maneira que livros seriam organizados em uma biblioteca de acordo com o assunto. Parte da habilidade de programar na linguagem LULL reside no conhecimento de onde encontrar certas funções, como procurar livros de viajem numa seção de viagens e não numa seção de ficção.

4.2.2.2 Programação em LULL

Escrever um programa em LULL, por mais complexo que seja, exige três passos - codificação, compilação e execução.

Codificação - É a estrutura e sintaxe de comandos com que se estrutura a linguagem em texto.

Compilação - É o processo no qual o código digitado numa determinada versão possibilita ao computador entende-lo para poder executa-lo.

Execução - É o processo no qual o código compilado é posto em funcionamento.

Qualquer linguagem de programação é feita de uma variedade de componentes, que coletivamente define a linguagem. Esses incluem comentários, variáveis, *if statements*, *loops*, funções de chamada e outros mais. A linguagem LULL não é exceção e assim como as demais possuí esses componentes que num conjunto final dão forma a linguagem.

Comentários

Comentários são utilizados em uma variedade de casos, tais como, criando cabeçalhos de arquivos, adicionando descrições ao código e comentando partes do

programa.

Na linguagem LULL os comentários são especificados no código do programa em forma de texto entre caracteres #. Qualquer texto entre os dois # é ignorado durante a compilação do código do programa.

Variáveis

Variáveis são valores que, como seu próprio nome implica, são permitidos variar. Por exemplo, uma variável como 'nome' pode ser requerido em um programa. Quando a variável é usada, o valor atual da variável assumida é substituído pelo novo valor da variável.

Declaração de variáveis

Antes de uma variável poder ser usada, ela deve ser declarada como sendo de certo tipo, sendo assim quando partes subseqüentes do programa vir a usar a variável, ela será usada corretamente.

Na linguagem LULL as variáveis podem ser de tipo simples ou complexo. Um tipo simples pode ser um real, inteiro, string, booleano ou objeto. Um tipo complexo é essencialmente uma estrutura de dado 'especial', tais como um *'display window'* ou uma geometria. As variáveis são declaradas com pequenas diferenças de acordo com o seu tipo.

If Statements

If statement são usados para avaliar quando certo critério é falso ou verdadeiro, e então subseqüentemente executar um comando. A linguagem LULL permite o uso de *if statement* simples, compostos ou *else if statement*.

Loops

Loops são usados para repetir seções particulares do código num número dado de vezes ou até certo critério ter sido atendido. Existem dois tipos de *loop* na linguagem LULL, chamado *while..do loop* e *do..while loop*. Eles diferem da localização na qual a expressão é checada. Não há algo igual a *for loop* como em outras linguagens de programação. Para efetivar um *for loop* na linguagem LULL é necessário incrementar

manualmente um contador.

4.2.3 Elaboração das rotinas de validação

Para a implementar as rotinas de validação no ambiente OO foi necessário realizar as seguintes atividades:

a. estudar as rotinas usadas no CAD;

b. estudar a linguagem de programação LULL;

c. analisar as funções existentes na biblioteca do *Gothic*;

d. estudar a modelagem de dados empregada no EGB;

e. escrever os códigos na linguagem LULL;

f. realizar testes aplicativos.

A seguir será abordado cada um dos itens mencionados acima, de maneira a esclarecer o que foi realizado em cada um.

Estudo das rotinas usadas no CAD

Esta parte do trabalho foi grandemente favorecida pelo fato de existir farta documentação das rotinas desenvolvidas para o uso nas linhas de produção da DSG. Do estudo das rotinas e sua estrutura foi possível extrair as principais funções que deveriam ser buscadas na biblioteca de funções do *Gothic*. As funções encontradas nestas rotinas são basicamente ferramentas pré-existentes no CAD tais como, seleção de elementos pelos seus atributos lógicos e gráficos (ferramenta *select by*) e alteração destes (ferramenta *modify*). Paralelamente a esta atividade foi analisada a estrutura lógica do programa (algoritmo) procurando entender como o mesmo funcionava no ambiente de CAD e se o mesmo seria adequado para ser empregado no ambiente orientado a objetos do *Gothic*.

Estudo da linguagem de programação LULL

O estudo da linguagem de programação LULL foi realizado por meio da análise do material bibliográfico existente na 1ª Divisão de Levantamento, basicamente manuais de programação e guias de iniciação rápida do software. Desta fase foi possível antever o grau de complexidade do sistema e também a potencialidade do mesmo, em termos de simulação das feições encontradas no mundo real. Como mencionada anteriormente a linguagem

LULL assemelha-se muito em termos de estrutura a antiga linguagem Pascal e a linguagem C. Dado o fato de que as rotinas MDL desenvolvidas para uso em ambiente de CAD são desenvolvidas em C não houve dificuldades maiores para entender seu funcionamento. Entretanto como a linguagem LULL trabalha sobre dados armazenados em banco de dados houve uma necessidade maior em compreender como o mesmo funcionava e como os conceitos da orientação a objetos eram aplicados sobre os dados geográficos.

Analise das funções existentes na biblioteca do *Gothic*

Como foi mencionado anteriormente o sucesso da programação no ambiente do *Gothic*, ou seja, das implementações das rotinas, reside basicamente no conhecimento das funções existentes no sistema. Neste ponto foi necessário o auxílio de operadores experiente na programação do sistema. Pelo intermédio destes profissionais foi possível se levantar a localização das funções mais adequadas para a execução das rotinas de validação. Neste ponto também ficou patente que sem o apoio destes profissionais o trabalho de implementação seria impossível, pois por si só o tempo necessário para completo entendimento do nível de complexidade da estrutura de programação do sistema inviabilizaria a confecção desta dissertação.

Estudo da Modelagem de dados do Espaço Geográfico Brasileiro

A fartura de material documentando a modelagem de dados do EGB disponível na 1ª DL bem como o fato daquela unidade ter participado ativamente da elaboração da mesma, facilitou consideravelmente o estudo realizado. Da análise da modelagem foi possível verificar como se enquadrava todas as categorias de validação, geradas naquela fase de produção em ambiente de CAD, dentro da definição de classes do sistema orientado a objetos. Também foi possível verificar detalhes importantes na definição das classes que poderiam ser utilizadas dentro da lógica de programação voltada para o sistema orientado a objetos. Por exemplo, foi possível atribuir às classes possuidoras de elementos lineares, tais como hidrografia, transportes dentre outras, a função de servirem como delimitadores de elementos de áreas.

Escrita dos códigos na linguagem LULL

A escrita da dos códigos das aplicações definidas para uso na validação dos dados foi realizada por um programador com experiência de uso na linguagem LULL. Basicamente foram escritas 25 rotinas, as quais possui embutidos em seu código fonte diversas funções retiradas da biblioteca do *Gothic*. O código fonte do aplicativo desenvolvido tem aproximadamente 10.000 linhas de programação e permitindo quando necessário à alteração de parâmetros e implementações adicionais. É importante salientar que boa parte dos procedimentos transcrita para o aplicativo final foi desenvolvida separadamente, como as ferramentas aplicadas na metodologia de validação dos dados no ambiente de CAD vistas anteriormente.

Testes com o aplicativo

Mesmo durante a escrita foram executados diversos testes com o aplicativo criado de maneira a depurar erros de sintaxe e erros de algoritmos. As rotinas integrantes do aplicativo de validação foram implementadas individualmente durante a fase de desenvolvimento, uma vez que o processo de criação de *softwares* praticamente induz a este procedimento. A ordem de aplicação das rotinas imposta pela metodologia e lógica do processo de validação, foi a responsável pela estruturação do aplicativo final. Mesmo considerando esses aspectos foram realizados dois testes preliminares com o material cartográfico oriundo da aquisição. Este material constituía-se de dois arquivos vetoriais oriundos do processo de restituição fotogramétrica que já haviam sido validados em ambiente de CAD. Os resultados apresentados no processo de validação foram bastante significativos tanto em termos de redução de tempo de produção como em qualidade final do produto.

4.3 Metodologia de Validação no ambiente OO

A metodologia desenvolvida para a realização da validação no sistema Gothic prevê a execução de três fases de processamento. A primeira fase realiza a validação dos objetos lineares e pontuais por meio da aplicação de rotinas que levam em conta algoritmos geométricos equivalentes aos encontrados em sistemas de CAD. A segunda fase aplica

sobre os objetos ferramentas baseadas em princípios topológicos modelados durante a definição das classes. A terceira fase, que também envolve a aplicação de regras topológicas, é caracterizada por tratar dos objetos de área. A seguir serão vistos com maior detalhes estas fases para que se possa ter uma visão mais abrangente sobre a metodologia e as rotinas desenvolvidas para validação no ambiente OO.

4.3.1 Fase I

Na primeira fase de validação a metodologia estabelecida faz com que os temas da TBCD sejam percorridos automaticamente um a um na seqüência estabelecida na Tabela 4.1. Durante este processo são executadas as rotinas sobre as classes de objetos pertencentes aos temas.

Tema da TBCD	Número de Ordem
Sistema transporte	1
Infra-estrutura	2
Edificações	3
Limites	4
Pontos de Referência	5
Hidrografia	6
Localidades	7
Atimetria (sem curvas de nível)	8
Altimetria (somente curvas de nível)	9
Vegetação	10
Extrativismo vegetal	11
Toponímia	12

Tabela 4.1 - Seqüência de temas da TBCD a serem percorridos na Fase I.

Os temas altimetria sem curva de nível e somente curvas de nível apresentadas na Tabela 4.1 são oriundos do desmembramento da categoria altimetria da TBCD. Outras duas mudanças em termos de classificação são: a reunião de todas as categorias de texto como se fosse um tema único da TBCD representado pela toponímia, e a separação de elementos oriundos do tema de vegetação num tema a parte denominado extrativismo vegetal. Esta

visão da TBCD pelo sistema foi estabelecida para flexibilizar a execução das rotinas desenvolvidas.

A seguir será dada uma descrição mais detalhada das rotinas desenvolvidas para esta primeira fase do processo de validação.

Rotinas da Fase I

Como mencionado anteriormente a primeira fase é responsável pela validação primária sobre os dados gráficos importados para dentro do BDOO. Para tanto foram desenvolvidas rotinas que levam em conta algoritmos geométricos equivalentes aos encontrados em sistemas de CAD. O processamento das rotinas prevê a execução em duas etapas, sendo a primeira de forma automática na qual o operador não interfere no processo e uma segunda na qual as rotinas maçam os problemas e o operador decide qual decisão deve tomar para corrigi-lo. A ordem de aplicação destas rotinas e suas descrições são dadas a seguir:

Etapa Automática

1º - EliminarDup_1:

Elimina objetos (lineares e pontuais), que se encontram na mesma classe. A rotina trabalha por meio de pesquisa feita numa área de influência (*buffer*) em torno dos objetos. A área de influência é definida de acordo com a escala do *dataset* no qual se está trabalhando. Por exemplo, para uma escala 1:25.000 é definida uma distância ao redor dos objetos de 2,0 m, que gera uma área de influencia na qual serão procurados os objetos duplicados. Caso seja identificado um objeto que se encontre completamente inserido nesta área de influência a rotina elimina-o do registro do banco de dados (Figura 4.1).

2º - FiltrarLinhas_1

Elimina vértices em excesso de objetos lineares por meio do uso do algoritmo Douglas Peucker (ver Capítulo 2). A distância máxima entre vértices usada como parâmetro, no algoritmo é estabelecida de acordo com a escala do *dataset* no qual se está

trabalhando. Por exemplo, para um *dataset* na escala 1:25.000 a distância adotada é de 1,5 metros.

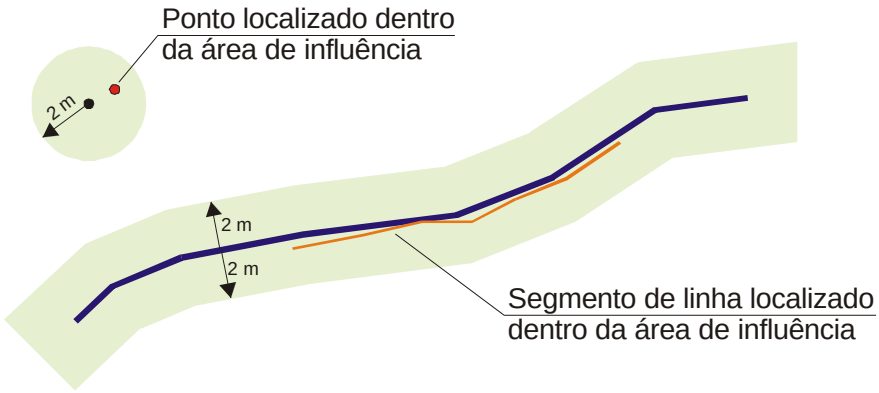


Figura 4.1 - Regra para eliminação de elementos adotada na rotina EliminarDup.

3º - AbrirLoops_1

Quebra os objetos lineares que se auto-interceptam. A quebra dos objetos criando vértices nos locais permite identificar geometrias que foram desenhadas de maneira equivocada, conforme pode ser visto na ilustração da Figura 4.2. Após a quebra do objeto os *loops* são isolados permitindo uma limpeza posterior por meio da aplicação da rotina EliminaDup_2.

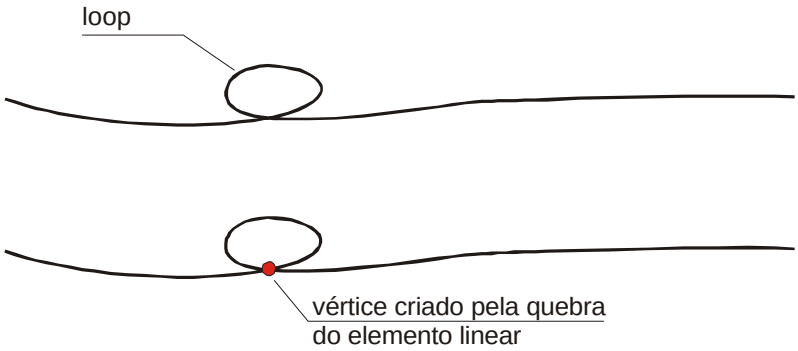


Figura 4.2 - Linha desenhada com *loops* antes e após a aplicação da rotina AbrirLoops.

4º - ExtLinhasFlut

Estende objetos lineares com vértices posicionados antes do alvo (*undershot*) até o encontro com outro objeto linear para que possa permitir a criação de uma interseção e futura definição de um nó. O valor máximo para estender o objeto linear até outro varia de

acordo com a escala do *dataset* que se está trabalhando e com a dimensão linear do objeto em questão. Por exemplo, para o caso de um *dataset* na escala 1:25.000, o valor limite para estender um elemento que possui mais que 30 m de comprimento é 15 metros Figura 4.3.

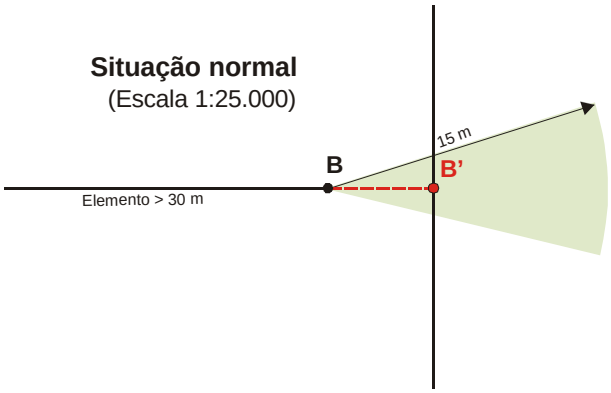


Figura 4.3 - Aplicação da rotina ExtLinhasFlut em *dataset* na escala 1:25.000.

Caso o objeto linear for menor que 30 m o valor máximo para estendê-lo é da metade do seu comprimento. Esta regra foi estabelecida para evitar a alteração da geometria dos objetos. Por exemplo, se existir um segmento de objeto linear com comprimento relativamente curto posicionado próximo a um outro objeto o algoritmo poderia se equivocar e reconstruí-lo a partir do vértice oposto ao desejado ocasionando uma alteração na geometria do objeto original. A Figura 4.4 ilustra esta situação e mostra que a regra lógica usada na construção da rotina que permite a manutenção da geometria de forma consistente.

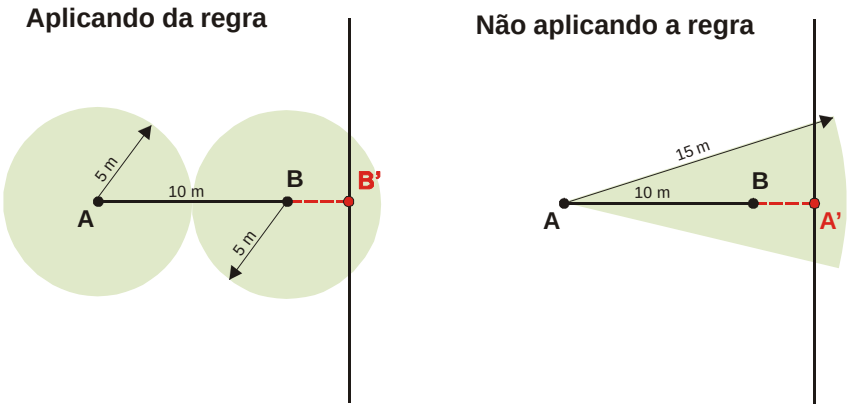


Figura 4.4 - Aplicação da rotina ExtLinhasFlut sobre elementos menores que 30 m.

5º - AbrirLoops_2

A rotina AbrirLoops_2 é idêntica a rotina AbrirLoops_1. Esta é aplicada novamente para eliminar *loops* dos objetos lineares que possivelmente podem ser gerados após a aplicação da rotina ExtLinhasFlut (Figura 4.5). Uma característica que não foi mencionada anteriormente e que deve ser destacada é que após a isolação dos *loops* o mesmo é dividido em duas partes para evitar erros durante o processamento das demais rotinas.

Figura 4.5 - Situação possível de alguns elementos lineares antes (a) e após a aplicação da rotina ExtLinFlut (b) e o *loop* gerado.

6º - GerarIntersecoes

A rotina gera interseções nos locais onde os objetos lineares se cruzam por meio da inserção de vértices quebrando-os. A execução desta rotina é feita visando a posterior eliminação de segmentos curtos por meio da aplicação da rotina DelLinhasPeq_1, bem como a possível geração de futuros nós (Figura 4.6).

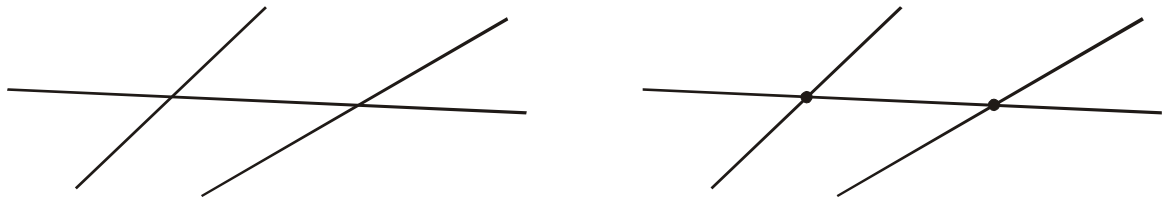


Figura 4.6 - Aplicação da rotina GerarIntersecoes em cruzamentos de elementos lineares.

7º - DelLinhasPeq_1

Elimina os segmentos curtos de objetos lineares gerados após a aplicação da rotina GerarIntersecoes. O segmento curto é caracterizado de acordo com o seu comprimento na escala do *dataset* que se está trabalhando. Por exemplo, para um *dataset* na

escala 1:25.000 a tolerância adotada para o comprimento do elemento é de 20 m. Entretanto se o segmento curto possuir um dos vértices coincidentes com o limite geográfico da folha (moldura) o mesmo não será eliminado. A implementação desta regra como condicionante na rotina previne a eliminação errônea de objetos que possam ter continuidade em folhas adjacentes. Caso o segmento curto mesmo contendo um dos vértices no limite geográfico da moldura possua comprimento menor que 7,5m será eliminado durante o processamento da rotina (Figura 4.7). Os demais segmentos curtos somente serão eliminados se:

a. possuir os dois vértices livres; ou,
b. um vértice livre e o outro conectado a dois ou mais objetos lineares.

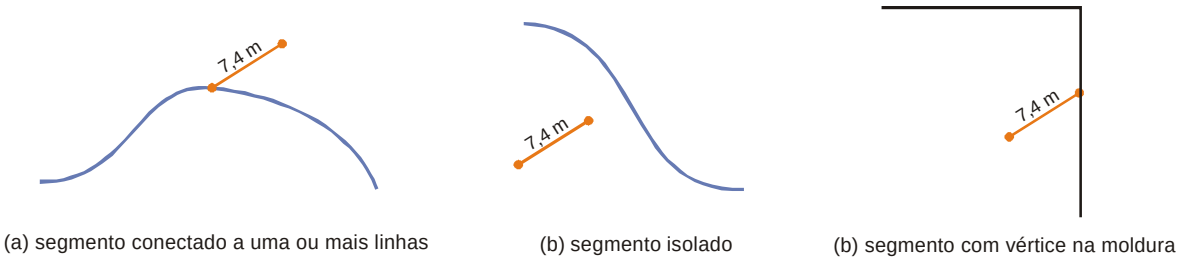


Figura 4.7 - Situação nas quais o segmento de linha curto será eliminado.

8º - EliminaDup_2

Elimina objetos duplicados após a aplicação das rotinas ExtLinhasFlut e GerarIntersecoes, por exemplo, situação ilustrada pela Figura 4.8. A rotina trabalha na mesma maneira que a EliminaDup_1 ou seja por meio da aplicação de uma área de influência ao longo dos objetos.

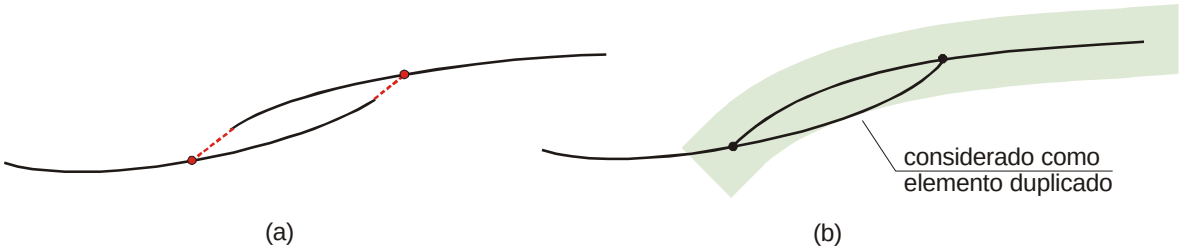


Figura 4.8 - Possível elemento duplicado: (a) gerado após aplicação das rotinas ExtLinhasFlut e GerarIntersecoes e (b) detectado para eliminação pela rotina EliminaDup.

9º - UnirLinhas_1

Une linhas contíguas de mesma classe e com os mesmos atributos.

10º - DelLinhasPeq_2

Igual a rotina DelLinhasPeq_1. A rotina é executada novamente para eliminar segmentos de linhas pequenos gerados pela aplicação da rotina EliminarDup_2. Por exemplo, dois segmentos de linhas curtos sobrepostos configuravam a inexistência de pontas livres, quando executado o EliminarDup_2 um dos elementos foi eliminado e configurou-se o segmento curto com ponta livre (Figura 4.9).

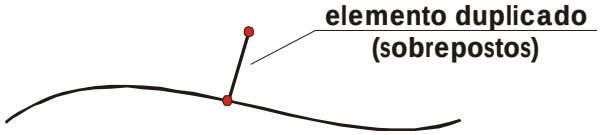


Figura 4.9 - Elementos sobrepostos após a execução de EliminaDup podem configurar segmentos curtos.

Outra finalidade da rotina é identificar e re-classificar segmentos curtos que possuam continuidade com elementos de outra classe, preparando-o para ser unido durante o processamento da rotina UnirLinhas_2.

11º - UnirLinhas_2

Idêntico ao UnirLinhas_1.

12º - FiltrarLinhas_2

Idêntico ao FiltrarLinhas_1. A rotina é executada novamente tendo em vista a união dos vértices criados pelas rotinas anteriores.

13º - ConectarCerca/ContVeg

A rotina estende o traçado das cercas e contorno de vegetação aos objetos lineares considerados delimitadores de áreas.

Etapa Iterativa

14º - FlagPta_Livre

Marca vértices livres existentes nos objetos lineares. A vértice livre é caracterizado após ter sido verificado se dentro de um raio de 2,0 mm na escala do *dataset* a partir deste for encontrado um outro objeto linear ou outro vértice livre. Os vértices livres que se encontrarem dentro desta área de influência são marcados e são considerados como possíveis vértices de ligação entre objetos lineares com os mesmos atributos (Figura 4.10). Caso seja encontrado dentro desta tolerância um vértice pertencente a outro objeto então os dois vértices livres são marcados para serem checados pelo operador.

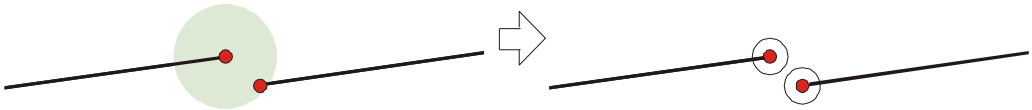


Figura 4.10 - Marcação dos vértices livres para serem checados pelo operador.

15º - FlagPeqSeg

Marca segmentos curtos de objetos lineares. O segmento curto é caracterizado pelas dimensões da Tabela 02, de acordo com a escala do *dataset* que se está trabalhando. A marca é posicionada sobre o segmento curto exatamente no ponto médio entre os vértices que o define.

Tolerância (m)	**Escala do *Dataset***
5	1:25.000
7	1:50.000
14	1:100.000
35	1:250.000

Tabela 4.2 - Tolerâncias da rotina FlagPeqSeg.

16º - FlagPeqAng

Marca vértices de objetos lineares que formam ângulos pequenos. O ângulo pequeno tem como limite máximo um valor de 15º (Figura 4.11).

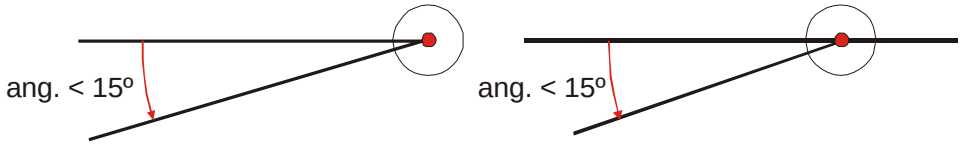


Figura 4.11 – Marcação de vértices que originam ângulos pequenos.

17º - FlagElemProx

Marca vértices que se encontram próximos a objetos adjacentes. Os valores limites para identificar a distância entre os vértices e os objetos é a mesma adotada na rotina FlagPeqSeg. Para o caso dos objetos lineares da classe altimetria (curvas de nível) é adotado um valor de 2 m na escala da carta para identificar os vértices.

18º - FlagInt/Loop

Marca vértices criados pela interseção de objetos lineares que configuram *loops* e interrupções. A Figura 4.12 ilustra algumas situações que podem ocorrer.

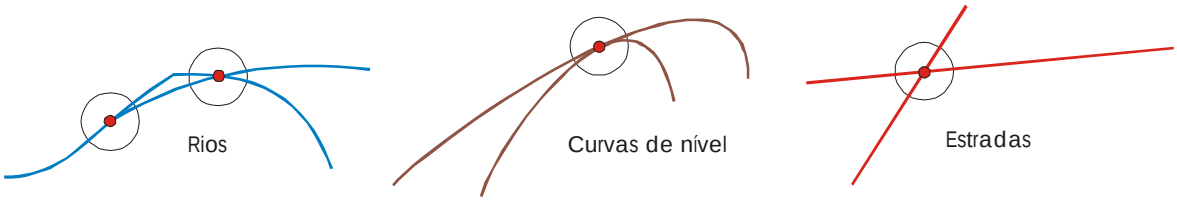


Figura 4.12 - Situação marcadas pela rotina FlagInt/Loop.

19º - FlagPtaLvContornos

Marca vértices livres pertencentes ao traçado de contornos de objetos de área (Figura 4.13).

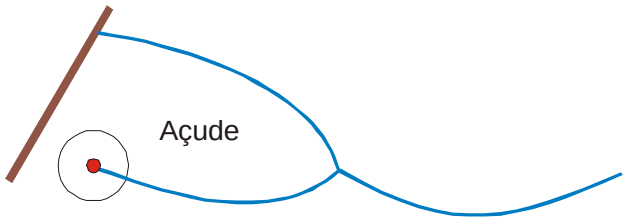


Figura 4.13 – Situação na qual o elemento de área está aberto e apresenta um vértice livre.

20º - Cadastra_altitude_cota

Atribui o valor numérico existente num objeto texto próximo ao objeto ponto cotado. A rotina é executada para todos os objetos ponto cotado desde que esteja contido dentro de uma área de influência em torno destes objetos texto com o valor a ser atribuído. Caso exista mais de um objeto texto próximo ao objeto ponto cotado a rotina marca este último para que o operador possa realizar a seleção do valor correto. Esta rotina foi desenvolvida para ser executada sobre os objetos ponto cotados que não possuem valores altimétricos cadastrados.

21º - Checa_altitude_cota

A rotina compara o valor do atributo altitude do objeto ponto cotado já cadastrado com o objeto texto mais próximo. Caso não haja igualdade entre os dois valores a rotina marca o objeto ponto cotado para que o operador realize a verificação.

4.3.2 Fase II

A segunda fase consiste na definição e aplicação das regras topológicas estabelecidas durante a implementação do processo de validação dentro do ambiente OO.

A definição consiste basicamente na ordenação de prioridades e relacionamentos entre classes de objetos para a aplicação das regras. Por exemplo, a classe de objetos rodovia pavimentada possui uma prioridade mais elevada em relação a classe arruamento, ou seja, no momento de decidir entre qual objeto deve ser modificado aquele que possuir prioridade maior terá preferência em termos de não alteração de sua geometria.

A aplicação das regras topológicas sobre as classes de objetos é realizada após a definição das prioridades. As regras topológicas criadas para o trabalho de validação de dados geográficos foram estabelecidas após a idealização prática de como os objetos do

Espaço Geográfico Brasileiro deveriam se comportar para atender as necessidades de criação de áreas que se aplicam neste modelo de dados.

Rotina da Fase II

Para execução da validação levando-se em conta as regras topológicas foi desenvolvida uma rotina que divide o processamento em duas partes.

A primeira consiste na aplicação das regras topológicas sobre a classe de objeto limite geográfico (moldura). O resultado da aplicação das topologias faz com que a geometria da classe mencionada se torne inalterável e assuma ações topológicas em relação às demais classes existentes. Os objetos sobre esta influência são ajustados de acordo com a geometria da moldura, ou seja, os objetos das classes que se encontrarem dentro do limite de influência topológica da moldura têm seus nós ou vértices estendidos até a moldura. O limite de influência usado como referência, para que os objetos tenham seus nós ou vértices capturados pelo limite geográfico de uma carta é de 50cm para uma carta na escala 1:25.000.

A segunda consiste na aplicação das regras topológicas sobre todos as demais classes de objetos. Aqui entra as prioridades entre classes de objetos estabelecidas nas regras de topologia da modelagem. O resultado da aplicação das topologias, neste caso, fará com que os objetos das classes se ajustem automaticamente, de acordo com as regras estabelecidas. Em outras palavras, os objetos das classes que possuírem prioridades menores em relação aos demais terão seus nós e ou vértices estendidos até o encontro ao traçado geométrico dos objetos de maior prioridade. O limite de influência referencial para que os objetos tenham seus vértices e ou nós estendidos é de 2m para um *dataset* na escala 1:25.000.

Tanto para a primeira parte como para a segunda parte, nos casos onde não foi possível o ajuste automático por meio da topologia ao marcados para serem verificados posteriormente pelo operador. Os primeiros testes da rotina da Fase II mostraram que as quantidades de erros apontadas após o processamento são baixas, da ordem de uma a duas dezenas de casos, por *dataset* de uma carta independente da escala.

4.3.3 Fase III

A terceira fase consiste na obtenção dos objetos de área a partir da utilização de parte da geometria de objetos lineares. Os objetos de área são formados a partir do contorno de segmentos de objetos lineares, de acordo com a Tabela 03. A validação na Fase III é dividida basicamente em três partes.

A primeira realiza o fechamento das áreas utilizando as classes de objetos considerados delimitadores, conforme relacionados na Tabela 4.3. A segunda classifica as áreas geradas, por meio de consultas ao atributo cor e aos centróides.

Áreas Tipo	Objetos delimitadores
Areas de infra-estrutura	limite_geográfico, contorno_de_infraestrutura
Edificacoes	limite_geografico, contorno_de_edificacao
Referencial	limite_geografico, contorno_de_referencial
Hidrografia	limite_geografico, contorno_de_hidrografia, barragem
Areas urbanas	contorno_de_edificacao, contorno_de_hidrografia, contorno_de_infraestrutura, contorno_de_referencial, contorno_de_localidade, contorno_de_relevo, contorno_de_vegetação, limite_geografico, acesso|arruamento, auto_estrada_div, auto_estrada_simp, barragem, caminho, canal, cerca, corte, ferrovia, linha_costeira, linha_drenagem, linha_marginal, muro, pista_competicao, quebra_mar, rod_n_pav_period, rod_n_pav_perm, rod_pavimentada, trilha, vala, elevado
Relevo	limite_geografico, contorno_de_relevo
Areas de vegetacao	limite_geográfico, contorno_de_edificação, contorno_de_referencial, contorno_de_hidrografia, contorno_de_infraestrutura, contorno_de_localidade, contorno_de_relevo, contorno_de_vegetacao, acesso, arruamento, auto_estrada_div, auto_estrada_simp, caminho, canal, cerca, corte, elevado, ferrovia, linha_costeira, linha_drenagem, linha_marginal, muro, pista_competição, barragem, quebra_mar, rampa, rod_n_pav_period, rod_n_pav_perm, rod_pavimentada, trilha, vala

Tabela 4.3 - Objetos delimitadores de área.

A terceira e última parte leva os objetos de área gerados e já classificados durante o processamento às classes de objetos aos quais pertencem, pois as áreas criadas são generalizadas em uma classe de objetos temporária.

Rotinas da Fase III

As rotinas criadas para realizarem os processos comentados acima empregam os segmentos de linhas dos objetos delimitadores de área para criar um conjunto de pequenas áreas, as quais não possuem qualquer identificação. Para que sejam identificadas corretamente as áreas são submetidas a um processo de classificação baseado em algoritmo geométrico. Este algoritmo realiza consultas ao: atributo cor do objeto oriundo da TBCD, para o caso do tema edificação e temas de hidrografia; e, centróide inserido durante o processo de aquisição de dados, para o caso das demais áreas.

É importante salientar aqui que os objetos de área embora estejam classificados pertencem a uma única classe genérica (de caráter temporário). Este procedimento se faz necessário por dois motivos a seguir expostos:

- verificação dos objetos de área que não foram classificados; e,
- necessidade de excluir áreas que representem "buracos" (*holes*) de áreas que foram construídas indevidamente.

No primeiro caso os objetos que apresentarem problemas de classificação são marcados e corrigidos posteriormente pelo operador. No segundo o processamento das rotinas detecta automaticamente áreas contidas em outras que se caracterizam como "buracos" (Figura 4.14).

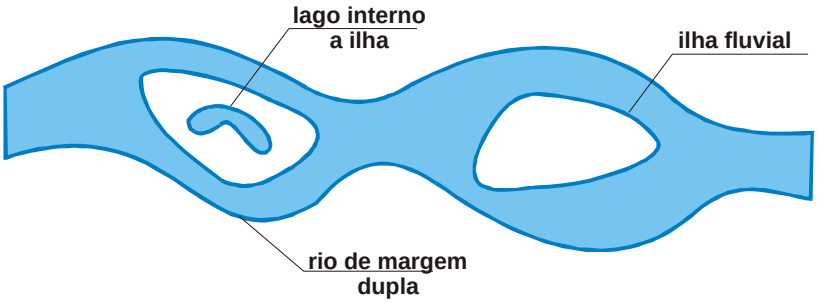


Figura 4.14 – Buracos existentes num objeto de área tipo rio de margem dupla.

Criação de objeto de área a partir do atributo cor

A classificação dos objetos de área realizada a partir da consulta do atributo cor é válida somente para os objetos oriundos dos temas da TBCD relativos a: infra-estrutura, transportes, edificações, relevo, referenciais e hidrografia.

A TBCD foi criada de maneira a não repetir dentro de uma mesma categoria o valor do atributo cor, que pode variar de 0 a 255. Considerando este fato o algoritmo realiza uma consulta ao atributo do objeto e verifica qual o valor numérico de cor está a ele associado. A identificação do objeto e conseqüente classificação são realizadas após tal constatação.

Criação de objeto de área a partir do centróide

A classificação dos objetos de área realizada a partir da consulta ao centróide é realizada para os objetos oriundos dos temas da TBCD relativos a: edificações e vegetação.

Durante a aquisição de dados são inseridos centróides no interior de espaços circundados por elementos lineares para posterior identificação e definição dos elementos de áreas. Os centróides são basicamente pontos que representam a origem de elementos de textos. Sua característica principal é possuir um texto por meio do qual se pode identificar áreas como pertencente a um determinado tipo de cobertura do solo. O algoritmo empregado usa esta informação contida no centróide para classificar a área na qual esta inserida. Para tanto ele realiza uma consulta tipo ponto contido em polígono e lê a informação texto do centróide classificando-o de acordo com o seu significado.

Para aprimorar a classificação de áreas genéricas de cobertura de solo e minimizar a iteração com o operador, foram criadas algumas condições.

Tendo em vista que as áreas oriundas dos objetos de hidrografia não são utilizadas durante o processamento, as áreas genéricas devem interagir com aquelas para que não haja inconsistências no espaço geográfico mapeado representado, ou seja, por exemplo, uma área de vegetação não pode ocupar o mesmo espaço que uma massa d´água. Para tanto após a criação dos polígonos estes são consultados para verificar tal regra, caso se confirme a sobreposição o polígono é eliminado.

Na seqüência serão procuradas sobreposições dos polígonos com áreas definidas como terreno exposto (minas, terreno rochoso, área urbana, areia, praia, dentre outras). Caso se confirme a sobreposição a área genérica é então classificada com tal.

Após isto são tratadas as áreas de vegetação, por meio de uma pesquisa que identifica o texto do centróide. Como comentado anteriormente, o centróide é implantado durante a aquisição do dado vetorial. Durante este procedimento pode ocorrer uma

generalização em relação ao posicionamento do centróide por parte do operador. Por exemplo, durante a inserção do centróide de vegetação o operador pode, levando em conta características do terreno, posicionar centróide da vegetação entre as feições, não individualizando as áreas para evitar a poluição do desenho cartográfico (Figura 4.15).

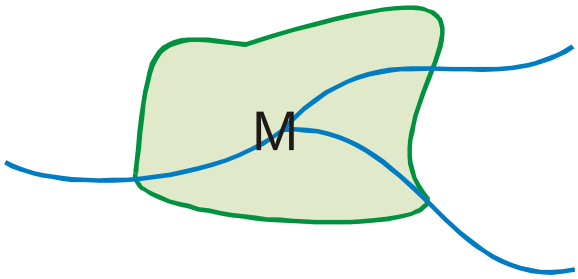


Figura 4.15 - Inserção de centróide de vegetação de forma generalizada sobre curso d´água.

Para evitar ter de inserir mais centróides para individualizar as áreas e resolver a classificação da área de maneira consistente foi estabelecido o que segue:

- 1º A área que possuir a origem do texto é classificada automaticamente conforme sua identificação;

- 2º É gerado um MBR (*Minimum Boundary Retangle* – Retângulo de Limite Mínimo) sobre o texto para verificar se o mesmo foi posicionado de maneira a englobar mais de uma área (situação ilustrada na Figura 4.15). As áreas que interceptarem o MBR são consultadas para constatar se já foram classificadas, aquelas que retornarem como não classificadas recebem a classificação do centróide do MBR; e,

- 3º As áreas não classificadas de acordo com o 1º procedimento ou o 2º, sofrem pesquisas de adjacência com áreas já classificadas.

O processo consiste em verificar o segmento de perímetro da área não classificada verificando se este segmento faz limite com uma área classificada. A Figura 4.16 ilustra dois casos típicos que podem ocorrer durante o processamento da rotina.

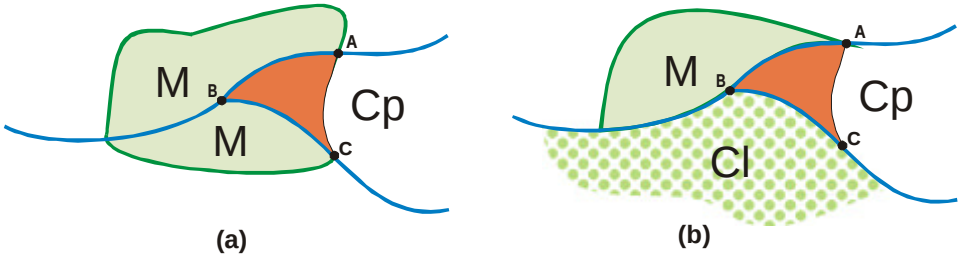


Figura 4.16 – Classificação de área não classificada por consulta de adjacência.

No caso da situação (a) da Figura 4.16 o processo confirma a existência de apenas um bom candidato para a definição da classe da área não classificada. O algoritmo executa esta definição como segue:

? Consulta-se o segmento AB. Se o mesmo faz parte do limite da área de mata (M), então a área concorre para ser classificada como M;

? Consulta-se o segmento AC. Se o mesmo faz parte do limite da área de campo (Cp), a área não poderá ser classificada nem como M nem como Cp;

? Consulta-se o segmento BC. Se o mesmo faz parte do limite da área de mata (M) então a área é classificada como M.

No caso da situação (b) da Figura 4.16 no caso da inexistência de bons candidatos a área deixa de ser classificada e é marcada para verificação por parte do operador.

Após a classificação de todas as áreas executa-se a Fase IV da validação conforme será abordada a seguir.

4.3.4 Fase VI

A quarta e última fase da validação consiste basicamente na definição e processamento da topologia de todos os objetos do *dataset* (tanto lineares como de áreas).

4.4 Avaliação da metodologia

Para a avaliação da metodologia de validação de dados em ambiente orientado a objetos foi estabelecido que seriam selecionados dois arquivos de cartas topográficas na escala 1:25.000. Os fatores que condicionaram a seleção do material a ser usado na avaliação foram basicamente três:

- As cartas deveriam ser consideradas “pesadas”, ou seja, com grande quantidade de elementos lineares e pontuais, quando comparadas com cartas adjacentes;

- Os dados vetoriais contidos nas cartas deveriam ser originários de restituição fotogramétrica, situação considerada mais abrangente para possíveis existências de problemas de traçado de feições; e,

- O material cartográfico selecionado deveria estar em condições de ser validado também em ambiente de CAD, ou já ter sido validado naquele ambiente.

As condições impostas por esses três fatores eram necessárias para que se pudesse realizar a avaliação dentro da ótica de se estar simulando uma situação real de produção de dados para SIG. Também tinha por finalidade possibilitar a análise comparativa entre as tecnologias orientada a objeto e a baseada em CAD, visando verificar as vantagens e deficiências de cada uma.

O material cartográfico selecionado para a avaliação foi obtido a partir da restituição fotogramétrica de duas cartas topográficas pertencentes ao projeto de mapeamento sistemático denominado Estabelecimentos de Ensino do Exército Brasileiro, localizado no estado do Paraná. A identificação das cartas de acordo com o número de Mapa Índice (MI) é 2867-2 e 2840-2. Estas cartas preenchiam adequadamente os pré-requisitos mencionados anteriormente sendo por isso selecionadas dentro do universo das cartas do mencionado projeto.

O primeiro passo realizado para realizar a avaliação foi a importação dos dados vetoriais para dentro dos *datasets* por meio do uso da ferramenta denominada FME (*Feature Mananger Engine* – Máquina de Gerenciamento de Feições). Esta ferramenta auxilia o *Gothic* nas operações de importação e exportação de dados, possuindo ainda algumas funcionalidades para o tratamento preliminar dos mesmos quando isto se faz necessário. Depois de realizada esta operação inicial seguiu-se à execução sobre os objetos do processo de validação.

4.4.1 Resultados Obtidos

Os resultados apresentados durante a execução parcial das fases de validação e após a aplicação completa da metodologia mostraram melhoras significativas em termos de redução de tempo de produção e na qualidade final dos dados.

Para se ter uma noção mais clara, por exemplo, da redução de tempo de produção, basta mencionar que na validação em ambiente de CAD de um único arquivo (daqueles selecionados para a avaliação) foram gastos cerca de 45 dias de trabalho. Já o mesmo trabalho no ambiente orientado a objetos usando a metodologia desenvolvida foi realizado em apenas 5 dias de trabalho.

O fator principal para atingir tal redução de tempo no processo de validação se deve basicamente aos aspectos mencionados a seguir:

- Algumas rotinas computacionais foram organizadas em grupos para a serem executadas de acordo com uma ordem lógica e de forma automática (rotinas da Etapa Automática da Fase I da metodologia de validação);

- A possibilidade de uso de regras topológicas no desenvolvimento das rotinas permitiu inserir certa inteligência nos dados tornando-os auto-ajustáveis a realidade do espaço geográfico modelado (Fase II da metodologia de validação);

- A criação de áreas aplicando o conceito de delimitadores e implementação de regras de consulta aos centróides e ao atributo cor dos objetos para obter a classificação possibilitou a verdadeira automatização para esta fase da validação (Fase II da metodologia de validação); e,

- A aplicação das regras de topologia sobre todos os elementos se mostrou extremamente eficiente permitindo um perfeito casamento entre os dados por meio do estabelecimento de prioridades de alterações entre os objetos (Fase IV da metodologia).

Uma observação importante que deve ser ressaltada, para entender o porque da diferença tão gritante em termos de tempo de validação, está relacionada na maneira como é efetuado o fechamento de feições de área nos dois ambientes. No ambiente de CAD o fechamento dos elementos de área é realizado por meio da aplicação de rotinas e da intervenção direta do operador durante todo o processo. Por exemplo, para a validação da categoria de vegetação da TBCD, que basicamente consiste do fechamento da maioria dos elementos de área existentes numa carta é o operador necessita executar 24 operações diferentes executando em cada uma processos de verificação e correção dos dados por meio da aplicação de ferramentas de edição existentes no CAD. Esta atividade corresponde à cerca de 2/3 do tempo total despendido na validação.

Com relação à qualidade final dos dados é importante lembrar que as regras topológicas aplicadas sobre os objetos armazenados no Banco de Dados Orientado a Objeto não permitem inconsistências. Este fator por si só já apresenta uma grande vantagem do ambiente orientado a objeto sobre o ambiente de CAD. Por exemplo, elementos de área para serem armazenados no BDOO devem estar obrigatoriamente fechados e classificados caso contrário seu armazenamento não é permitido. No ambiente de CAD, por sua vez, não é incomum encontrar elementos de área abertos ou de linhas não ligados ocasionando inconsistentes em relação aos demais. Os resultados, portanto para a constatação da

qualidade dos dados, após a aplicação da metodologia de validação, levou em conta esta consideração.

4.4.2 Considerações Finais

Durante a realização dos estudos visando a implementação de uma nova metodologia de validação para os dados surgiram diversas oportunidades nas quais foi possível compreender o potencial do sistema para emprego nas atividades de produção cartográfica. A diferença encontrada entre os dois ambientes estudados (CAD x OO), por sua vez, esclareceu as principais dúvidas existentes em relação ao sistema *Gothic*. As vantagens deste em relação ao CAD está basicamente na maneira de pensar os dados. Enquanto o sistema CAD visualiza entidades ponto, linha e polígono apenas como feições estáticas dentro de um espaço geográfico, o sistema OO enxerga as entidades como objetos possuidores de comportamentos e sujeitos a restrições topológicas.

O sistema *Gothic* em relação ao sistema CAD apresenta as seguintes vantagens:

- Permite que mais de um operador trabalhe sobre o mesmo espaço geográfico (*dataset*);
- Permite o desenvolvimento de rotinas de automação de processos;
- Garante a integridade dos dados pela aplicação de restrições topológicas sobre os dados trabalhados;
- Permite realizar consultas espaciais sobre os dados geográficos armazenados;
- Permite a simbolização cartográfica dos dados geográficos armazenados de forma automática;
- Possibilita o uso de topologias para elaboração de restrições espaciais sobre os dados.

Em contrapartida o sistema apresenta as seguintes desvantagens para uso na produção cartográfica:

- Necessita de programadores experientes, com profundo conhecimento das funções existentes no sistema;
- Trabalha com uma interface gráfica não conhecida e com recursos de produtividade limitados;

- Roda em ambiente windows por meio de programas que simulam o ambiente para o qual foi desenvolvido (*emulador*), a saber UNIX;
- O acesso ao sistema se faz por meio de redes locais aos servidores de BD, podendo ocorrer congestionamento de transações ao se processar grande quantidade de dados.

Um dentre os diversos aspectos positivos apresentados pelo ambiente orientado a objetos em oposição ao ambiente de CAD é que os dados por estarem armazenados no BDOO podem ser editados automaticamente e representados com a simbologia cartográfica conforme preconizado no manual de representações cartográficas T34-700.

Esse fato permite uma mudança radical em termos de linha de produção permitindo fundir duas fases distintas da produção cartográfica em uma única fase. O fluxo de produção existente hoje na 1ª Divisões de Levantamento apresenta a distribuição das fases de produção de acordo com a Figura 4.17. Dentro daquele contexto o Tratamento de Dados é constituído pela validação que fornece dados para a edição vetorial e para o BDOO que os emprega na edição automática. Se a validação passa a ser executada dentro do BDOO é possível então realizar a alteração do fluxo de produção por meio da criação de uma fase a partir da fusão das fases de validação e edição automática. A Figura 6.18 ilustra esta proposta de alteração por meio da criação de uma fase denominada Tratamento de Dados.

Cabe mencionar que a edição vetorial realizada em ambiente de CAD seria completamente substituída pela edição automática, deixando de existir no fluxo de produção. Embora esta decisão possa parecer um pouco radical é importante lembrar que a edição vetorial já não vem sendo realizada desde que os dados passaram a ser inseridos no BDOO. O motivo principal para a adoção de tal medida pode ser justificado por motivos de ordem técnica e econômica, pois o tempo destinado para editar uma única carta de maneira vetorial em ambiente de CAD de acordo com os procedimentos usuais quando comparado com a edição automática é da ordem de 1:5, ou seja, com o tempo gasto para editar uma carta em ambiente de CAD pode-se editar 5 no ambiente orientado a objetos.

A alteração da linha de produção possibilitaria ainda a redução do número de operadores alocados para a função de validação permitindo uma redistribuição de operadores para outras fases da linha de produção, por exemplo, aquisição de dados via

vetorização ou atualização via imagens orbitais. Dentro do contexto de otimização de processos o custo benefício positivo para a DSG seria considerável, tendo em vista que ainda existe boa parte do material cartográfico em meio analógico, que deve ser convertido para o meio digital e atualizado.

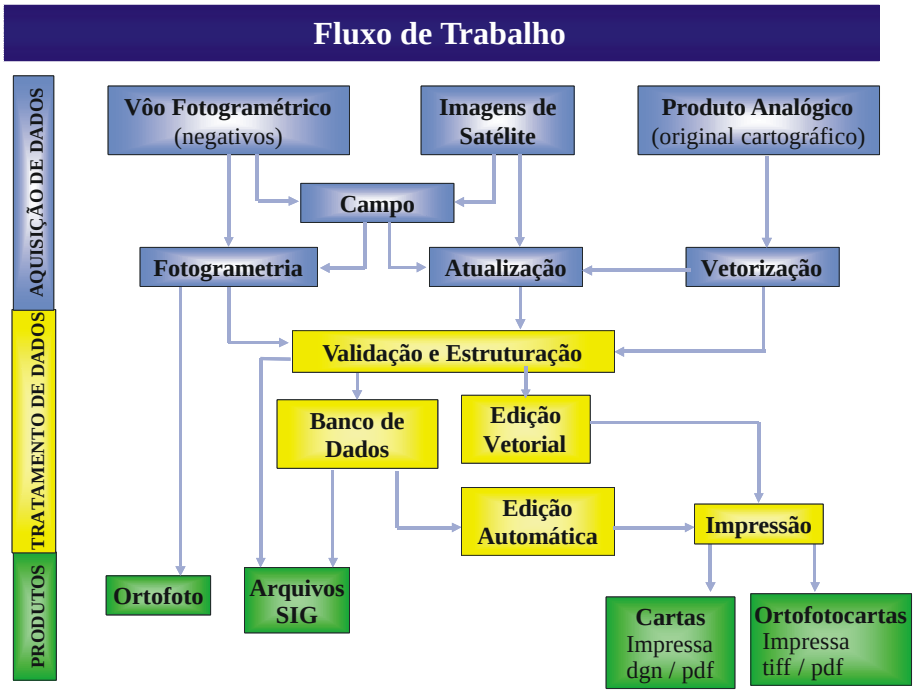


Figura 4.17 - Fluxo de trabalho atual na 1ª DL.

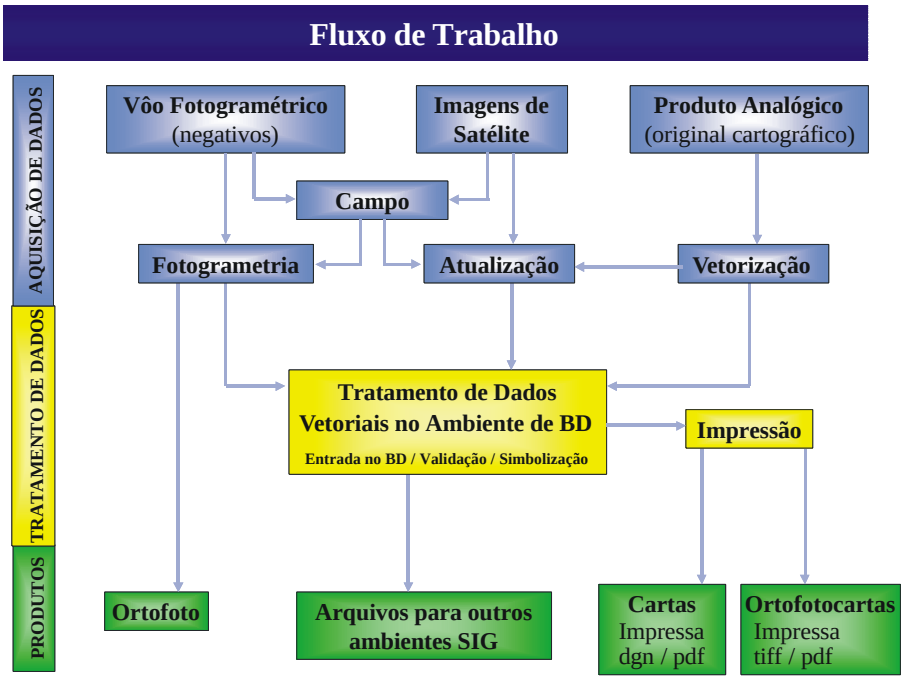


Figura 4.18 - Fluxo de trabalho proposto

A possibilidade de realizar a aquisição de dados dentro do sistema *Gothic* é extremamente interessante, pois a partir da implementação das fases de validação e simbolização cartográfica (edição) poder-se-ia pensar numa linha de produção completamente diferente da existente atualmente. O fato de se poder realizar a aquisição dos dados num ambiente onde as regras topológicas estivessem ativas em tempo real (*on-the-fly*) diminuiria muito as inconsistências dos dados melhorando consideravelmente a qualidade dos mesmos. Resultaria também em uma redução do tempo de validação, a qual seria executada, de certa forma, paralelamente a aquisição. Uma linha de produção nestes moldes seria algo revolucionário para automatização cartográfica, pois aquisição tratamento e simbolização deixariam de estar seqüencialmente ligados e passariam a fazer parte de um conjunto único de processos.

O domínio da tecnologia orientada a objetos, usada em de dois terços da linha de produção cartográfica é um bom começo para futuras perspectivas de automação de processos no âmbito da cartografia automatizada.

Capítulo 5 – CONCLUSÕES E RECOMENDAÇÕES

5.1 Conclusões

O aprimoramento atual de metodologias e técnicas desenvolvidas exclusivamente para a produção em ambiente de CAD tem permitido que a otimização de processos vise o aumento do índice de produção, via diminuição do tempo de confecção, e a melhoria da qualidade dos produtos cartográficos. Entretanto, apesar de todos esses esforços voltados para este ambiente os resultados obtidos correspondem apenas a uma parcela daquilo que seria necessário para suprir a demanda instaurada nos últimos anos.

Este trabalho de dissertação inserido dentro deste contexto realizou estudos visando à elaboração de procedimentos de validação dentro do sistema *Gothic*, orientado a objeto, com o objetivo de automatizar processos de validação geométrica, com o intuito de aumentar o ganho produtivo sem perda de qualidade. Para tanto foram abordados alguns aspectos teóricos e práticos relacionados à produção cartográfica na DSG.

Como se pode notar durante o trabalho, foram abordados no Capítulo 2 os principais conceitos relacionados à cartografia em meio digital e aos SIG, bem como, apresentada uma das metodologias de aquisição, na qual foram tratados alguns aspectos relacionados às principais fontes de erros. Ainda no Capítulo 2, foram apresentados os principais conceitos da tecnologia orientada a objetos e as principais características do sistema *Gothic* adotado pela DSG. No Capítulo 3 foi realizada uma descrição pormenorizada da metodologia de validação de dados em ambiente de CAD adotada atualmente pela DSG nas suas linhas de produção. E por fim foram apresentados os resultados do trabalho de implementação de nova metodologia de validação de dados em ambiente orientado a objetos, bem como foi realizada uma descrição das rotinas desenvolvidas no Capítulo 4.

Chegou-se à conclusão de que os ambientes de CAD e Orientado a Objetos, embora não pareçam num primeiro momento, são muito diferentes, e que esta tecnologia, embora tenha surgido a alguns anos atrás, ainda constitui-se num campo vasto para

pesquisa, envolvendo diversos conceitos que podem ser aplicados na automatização de processos da produção cartográfica.

A possibilidade de tratar os dados adequadamente para uso em SIG dentro de um sistema de banco de banco de dados baseado em ambiente orientado a objetos apresentou vantagens impares em relação ao ambiente de trabalho convencional (CAD) adotado pela DSG. Estas vantagens repercutiram não somente na qualidade final do produto gerado, mas também nas possibilidades de manuseio do mesmo num ambiente propicio ao armazenamento e análise espaciais.

Foram desenvolvidos diversas rotinas computacionais e estruturados novos procedimentos operacionais para a validação geométrica dos dados gerados na aquisição. A sedimentação dos procedimentos englobando numa estruturação lógica as rotinas permitiu a elaboração de uma metodologia preliminar que pode servir de ponto de partida para uma metodologia definitiva.

Com relação aos testes realizados com a metodologia elaborada, pode-se concluir que o objetivo foi atingido, pois após a realização obteve-se a diminuição do tempo de produção significativa sem a perda de qualidade. Os testes mostraram que o tempo de validação das cartas ficou sendo da ordem de 30 horas para uma carta na escala 1:50.000 oriunda de restituição. Enquanto que a qualidade pode ser verificada por meio das restrições impostas aos dados pelo sistema de armazenamento baseado em ambiente orientado a objeto.

Verificou-se também durante os primeiros testes a necessidade do desenvolvimento de novas rotinas ou o aprimoramento das já desenvolvidas, para adequar a metodologia a situações não previstas. É importante frisar que por estar ainda em processo de testes as rotinas criadas ainda devem ser encaradas como não definitivas, portanto passíveis de depurações e alterações visando seu aprimoramento.

Este trabalho serviu também para despertar o interesse para o uso da tecnologia orientada a objetos em pesquisas visando à implantação de uma linha de produção mais enxuta, com menos fases de trabalho e com um maior controle de qualidade para os produtos cartográficos em meio digital, hoje disponibilizados pela DSG.

5.2 Recomendações

As possibilidades de pesquisa em fases de aquisição de dados dentro do sistema visando futuras implementações é algo que deve ser encarado com seriedade e deve merecer atenção especial, pois o ambiente estudado mostrou estar perfeitamente capacitado para atender as expectativas de tais atividades. Portanto seria extremamente interessante a realização de pesquisas visando a integração da vetorização ao que já foi desenvolvido até agora.

Outra faceta que não foi explorada neste trabalho é a possibilidade de abordar de maneira adequada o problema oriundo da generalização cartográfica. Com relação a este assunto o sistema mostrou ser capaz de possibilitar a criação de soluções inteligentes que merecem ser abordadas em trabalhos futuros. É possível que com iniciativas voltadas a pesquisa de automação nesta área possa se chegar a resultados surpreendentes.

Para a utilização da metodologia sugerida na linha de produção da DSG deve-se ter em mente que ainda é necessário realizar ainda alguns testes com cartas oriundas da vetorização e fora dos padrões estabelecidos na TBCD. Estes testes devem ser realizados com o intuito de aprimorar as ferramentas desenvolvidas para uma melhor adequação das mesmas, aos diversos aspectos inerentes a dados adquiridos via CAD.

Outro aspecto importante a respeito do desenvolvimento das rotinas aplicadas na metodologia de validação e que merece atenção especial é que ainda existem funções do sistema *Gothic* que ainda não foram exploradas e que possivelmente podem ser úteis para o aprimoramento da metodologia sugerida neste trabalho. Por fim cabe dizer que as funções do *Gothic* usadas no desenvolvimento das rotinas e as regras topológicas aplicadas correspondem a uma pequena parcela daquelas existentes em sua biblioteca de funções e os recursos de elaboração de regras dentro do sistema.

REFERÊNCIAS BIBLIOGRÁFICAS

ALDRED, B. K.. **Point-in-polygon algorithms**. IBM, Peterlee, UK, 1972.

ALVES, D. S. **Sistemas de Informação Geográfica.** In: SIMPÓSIO BRASILEIRO DE GEOPROCESSAMENTO, 1., 1990, São Paulo. Anais. São Paulo:USP, 1990.

ABLER, Ronald F. **"Awards, Rewards, and Excellence: Keeping Geography Alive and Well."** The professional Geographer, 1988.

ALVES, D. S. **Modelo de Dados para Sistemas de Informações Geográficas.** São José dos Campos, SP: USP, 1994. Tese de Doutorado.

ANDERSON, James R.. **Land use classification schemes used in selected recent geographic applications of remote sensing: Photogramm**. Eng., 1971

ANDERSON, James R.; HARDY, Ernest E.; and ROACH, John T., 1971. **A land-use classification system for use** with remote-sensor data. U.S. Geol, Survey Circ., 1971.

ANTENUCCI, J.C.; BROWN, K.; CROSWELL, P.L.; et al. **Geographic Information Systems: A guide to the technology**. Van Nostrand Reinhold, 1991.

ARONOF, S. **Geographic Information Systems: a management perspective**. Canada: WDL Publications, 1989.

AUTODESK. **AutoCAD 2000 Manual Técnico**. Autodesk, 2000.

BEARD, K. **Theory of the Cartograpic Line Revisited: Implications for Automated Generalization**. Cortographica, 1991.

BICKMORE, D. **A database for a botanical plant collection.** In Computer-aided landscape desing: principles and pratice. Ed. B. M. Evans, Scotland, 1984.

BURROUGH, P.A. **Principles of Geographical Information Systems for Land**

Resources Assessment. Oxford: Clarendon Press, 1986.

CÂMARA, Gilberto. **Análise de Arquiteturas para Bancos de Dados Geográficos Orientados a Objetos.** São José dos Campos, SP: USP, 1994. Tese de Doutorado.

CÂMARA, G.; CASANOVA, M. A.; HEMERLY; A.; MEDEIROS, C.M.B.; MAGALHAES, G. **Anatomia de Sistemas de Informação Geográfica.** X Escola de Computação, Campinas: SBC, julho 1996.

DENT, B. D. **Principles of thematic map design**. Readings MA: Addison Wesley, 1999.

DEMERS, Michael N. **Fundamentals of Geographic Information Systems**. New York:Wiley, 1997.

DIBIASE, D. **Visualization in earth sciences.** Earth and Mineral Sciences. The Pennsylvania State University, 1990. Bulletin of the college of Earth and Mineral Sciences.

DOUGLAS, D. H. and PEUCKER, T. K. **Algorithms for the Redution of the Number of Points Required to Represent a Line or its Caricature**. The Canadian Cartographer, 1973.

DSG. **Normas Provisórias para Validação de Dados – NPVAD**. Diretoria de Serviço Geográfico, 2000.

DSG. **Manual de Vetorização**. Diretoria de Serviço Geográfico, 1999.

DSG. **Tabela da Base Cartografia Digital - TBCD**. Diretoria de Serviço Geográfico, 2001.

DSG. **T34-700 Manual Técnico: Convenções Cartográficas, 1ª e 2ª Partes**. DSG, 2002.

EASTMAN, J. R. **IDRISI for Windows User´s Guide**. Worcester, MA: Clark Labs for Geographic Technology an Geographic Analysis, 1997.

ESRI, ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. **Understanding GIS: The ARC/INFO Method**. Redlands, CA: Environmental Systems Research Institute, 1991.

ESRI. **ArcGIS User´s Guide**. ESRI, 2002.

ESRI. **Digital chart of the World for use with Arc/Info software**. Redlands: ESRI. 1993.

GEMAEL, Camil. **Ajustamento de Observações**. UFPR Editora, 2000.

GEOGRAPHER´S CRAFT. Disponível: http//www.utexas.edu/depts/grg/gcraft/notes/intro/intr.html#Geog

GOODCHILD, Michael F. **Geographical data modeling**. In: FRANK, A. U.; GOODCHILD, M.F. Two Perspectives on Geographical Data Modelling**.** Santa Barbara, CA: National Center for Geographic Information & Analysis/NCGIA, 1990. (Technical Paper 90-11)

GOODCHILD, Michael F. **Integrating GIS and environmental modeling at global scales**. In: GIS/LIS '91. Proceedings**.** Atlanta, 1991. V.1.

GOODCHILD, Michael F.. **Geographical data modeling.** Computers & Geosciences**,** London, v.18, n.4, p.401-408, 1992.

INPE. **Tutorial de Geoprocesamento**. INPE, 2004. Disponivel: http://www.dpi.inpe.br/spring/portugues/manuais.htmla

INPE. **Manual do SPRING**. INPE, 2004. Disponivel: http://www.dpi.inpe.br/spring/portugues/manuais.htmla

INTERGRAPH. **Modular GIS Enviroment: User´s Guide**. Integraph, 2001.

INTERGRAPH. **MicroStation User´s Guide**. Intergraph, 2001.

INTERGRAPH. **MicroStation Development Language Couse Guide**. Intergraph, 1991.

LASER-SCAN. **Gothic System Management**. Training Course. Laser-Scan, Cambridge, 2000.

LASER-SCAN. **Managing and Configuring the Gothic Environment**. User´s Guide. Laser-Scan, Cambridge, 2000.

LASER-SCAN. **The Pratical Guide to Database Concepts v1.0**. Training Course. Laser-Scan, Cambridge, 2000.

LASER-SCAN. **LAMPS2 Operations**. Training Course. Laser-Scan, Cambridge, 2000.

LASER-SCAN. **Gothic Database Concepts v1.1**. Training Course. Laser-Scan, Cambridge, 2000.

LASER-SCAN. **Advanced Programming in LULL**. User´s Guide. Laser-Scan, Cambridge, 2000.

LAURINI, Robert; THOMPSON, Derek. **Fundamentals of Spatial Information Systems.** San Diego: Academic Press, 1992.

MACEACHREN, A. M; TAYLOR, D. R. F.. **Visualization in modern cartography**. Modern cartography series, Pergamon Press, Oxford, 1994.

MACDOUGAL, E. B.. **The Accuracy of the overlays**. Landsc. Plan, 1975.

MAGUIRE, D.J.; GOODCHILD, M.F.; RHIND, D. (Eds.) **Geographical Information Systems: Principles and Applications.** 2 Vol., Longman Scientific & Technical, 1991.

MAGUIRE, David J. **The raster GIS design model - a profile of ERDAS**. Computers and Geosciences, v.18, n.4, p.463-470, 1992.

NATIONAL CENTER FOR GEOGRAPHIC INFORMATION AND ANALYSIS **NCGIA Core Curriculum.** Edited by Michael F. Goodchild and Karen K Kemp, Santa Barbara University of California, 1990.

OLIVEIRA, Cêurio de . **Curso de Cartografia Moderna.** 2. ed., Rio de Janeiro: IBGE. 1993.

PAREDES, E. A. **Sistema de Informação Geográfica - Principios** e **Aplicações: Geoprocessamento,** São Paulo: Ed. Erica. 1994.

PERKAL, J.. **On the length of empirical curves**. Discussion paper 10, Ann Arbor, Michigan Inter-University Community of Matematical Geographers, 1966.

PERKAL, J. **An Attempt at Objective Generalization**. Michigan Inter-University Community of Matematical Geographers, Discussion Paper No 10, 1996.

PEUQUET, Donna J. **A conceptual framework and comparison of spatial data models**. **Cartographica,** 1984.

RAIZ, Erwin. **Cartografia Geral (Tradução).** Editora Científica, RJ, 1969.

SAUNDERS, C. A. Barreto. **Notas de Cartografia**. IME, RJ, 1982. Apostila.

SMEDLEY, B. And ALDRED, B. K.. **Problem with geodata**. In Data base tchniques for pictorial applications (ed. A. Blases). Springer-Verlag, Berlin, 1980.

TAYLOR, D. R. Fraser. **Geographic Information Systems: An Introdution**. Pergamon Prees, Oxford, England, 1991.

TEIXEIRA, A. et. al. **Qual a melhor definição de SIG**. Revista FATOR GIS, n.11, 1995.

TIMES, Valéria C. **MGeo: Um Modelo Orientado a Objetos para Aplicações Geográficas.** Recife-PE: UFPE, 1994. Dissertação de Mestrado.

VIEIRA, Marina Teresa Pires. **Banco de Dados Orientado a Objetos**. Universidade Federal de São Carlos Departamento de Computação, apostila, 2001.

KRAAK, Menno; ORMELING, Ferjan. **Cartography: Visualization of Spacial Data**. Longman, London, England, 1996.

WOLF, Paul R.; DEWITT, Bon A. **Elements of Photogrammetry: With Applications in Gis**. McGraw-Hill College, 2000.